Prediction of Driving Ability in Healthy Older Adults and Adults with Alzheimer's Dementia or Mild Cognitive Impairment

Petra Ann Hoggarth

A thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy

in the University of Canterbury, Christchurch, New Zealand

August 2011

Abstract

Normal ageing is associated with decline in visual, cognitive, and physical functioning, with concurrent increases in the incidence of chronic medical conditions, including cognitive disorders. Determining when age-related changes have adversely affected a person's ability to drive safely is a complex task, particularly when cognitive disorders such as mild cognitive impairment and dementia are present.

The aim of this research was to assess the utility of a number of off-road measures in predicting Pass and Fail outcomes for older drivers on a blinded on-road driving assessment with a driving specialist occupational therapist and a driving instructor, which is considered the 'gold standard' measure of driving ability. The off-road measures included standardized cognitive tests, computerized sensory-motor & cognitive tests, medical conditions, and personality measures. The research project comprised three studies.

In Study 1 (Healthy Older Drivers study), 60 drivers with no diagnosed cognitive disorder ('cognitively-unimpaired'), aged 70-84 years (mean age 76.7, 50% male), completed standard cognitive tests, computerized sensory-motor and cognitive tests (*SMCTests™)*, and measures of personality. Results were used to form classification models for on-road assessment Pass and Fail outcome. Sixteen participants failed the on-road assessment. A backwards stepwise binary logistic regression model selected a measure of executive function and a computerized measure of visuomotor planning and coordination as the best predictors. Following leave-one-out cross-validation, this model was estimated to correctly predict 60% of an independent group of cognitively-unimpaired older drivers into on-road Pass and Fail groups.

In Study 2 (Healthy Driver Follow-up study), 56 participants from the Healthy Older Drivers study were followed for 24 months using annual telephone interviews to assess driving behaviour, driving attitudes, medical conditions, and the occurrence of crashes and receipt of traffic offences. Official data regarding crashes and traffic offences were also obtained. The aim was to determine whether either the on-road Pass/Fail classification or the off-road measures could predict subsequent crashes and offences. Failing the on-road assessment was not associated with higher crash or offence rates and there were only two baseline measures that predicted crashes or offences (i.e., distance driven at baseline testing and, paradoxically, a lower error score on a measure of visuomotor planning and coordination). However, drivers

who reported more distress associated with their medical condition(s) were more likely to have had a crash or offence at 24 months. The outcomes of the Healthy Older Drivers and Healthy Driver Follow-up studies suggest that there is little value in off-road or on-road assessment of cognitively-unimpaired older drivers due to the weak relationship with future negative driving outcomes. However, distress associated with medical conditions may be a useful measure.

Study 3 (Dementia and Driving study) recruited a sample of 60 driving assessment centre referrals with mild cognitive impairment or Alzheimer’s dementia. These participants, aged 58-92 years (mean age 77.9, 60% male), performed a computerized battery of sensory-motor and cognitive tests and a formal blinded on-road driving assessment. A backwards stepwise binary logistic regression model selected measures of reaction time and movement speed of the upper limbs, visuomotor planning and coordination, and sustained attention. Following leave-one-out cross-validation, this model was estimated to correctly predict 68% of an independent group of drivers with mild cognitive impairment and Alzheimer’s dementia into on-road Pass and Fail groups. A subsample of 32 participants completed additional standard cognitive tests and provided information on medical conditions. A binary logistic regression model in this subsample was formed which selected measures of verbal fluency, the presence of heart disease, and a comprehensive cognitive screen. Following leave-one-out cross-validation, this model would be expected to correctly classify 75% of an independent group of drivers with mild cognitive impairment and Alzheimer’s dementia into on-road Pass and Fail groups. The three measures in this model could be performed in around 35-50 min in a primary health setting.

It is concluded that off-road and on-road assessment of older drivers with no diagnosis of cognitive or neurological disorder is an inaccurate and inefficient use of driving assessment resources, both for the prediction of on-road driving performance and for predicting future crashes and traffic offences. The Dementia and Driving study found a model comprising three measures that could be performed in a primary health setting with reasonable accuracy for correctly classifying people with mild cognitive impairment and Alzheimer’s dementia who go on to Pass and Fail an on-road driving assessment.

At 12:51pm on Tuesday the 22nd of February, while I was in the in the final stages of preparing this thesis, the city of Christchurch and town of Lyttelton were struck by a devastating earthquake. While myself and my colleagues escaped from the Van der Veer Institute building unharmed, scores of people in the central city were killed and many more seriously injured. Large parts of our beautiful historic central city and Lyttelton have been destroyed.

This thesis is dedicated to the people of Christchurch and Lyttelton. We are resilient, our city will be rebuilt, and we will endure.

Acknowledgements

I would like to thank several people who, through their assistance and friendship, helped make this research project possible.

Firstly, I would like to thank my supervisors, Associate Professor Richard Jones, Dr Carrie Innes and Associate Professor John Dalrymple-Alford. Richard was unwavering in his support for the project, always made time in his busy schedule to meet and discuss the project, and had an inquisitive mind which helped to shape project goals and interpret results. Carrie was always available to assist in details of project design, data collection, analysis of data, and moral support. Her expertise in the field of driving assessment was of immense value and her day-to-day interest and support were invaluable. John kept me on my toes with suggestions for refining the project and its interpretation. His detailed reviewing of written output was always helpful and I always knew that John would bring a fresh perspective to the task. His expertise in neuropsychological testing was invaluable for this project.

I must also thank the other students and staff at the Van der Veer Institute with special mention to my cubicle-mate of four years, Phoebe Macrae, whose sense of humour helped even the darkest times to feel better. Many a morning tea was had with students and staff, and these were indeed required to provide a welcome break from test scoring, data entry, and statistical analysis. I would like to thank fellow PhD students Hannah Farr, Saskia van Stockum, Charlotte Graham, Tracy Melzer, Sarah Wright, and Oshrat Sella for their camaraderie – something we all need when undertaking such a research project.

I would like to thank the approximately 120 participants in my study whose generosity with their time was so appreciated. The many cups of tea and biscuits and the many recitations of life stories along the way made the data collection process more bearable for both parties.

I am very grateful for funding I have received at various points during the course of the project. A Masters Scholarship from the University of Canterbury and a Christchurch Neurotechnology Scholarship awarded by the Christchurch Neurotechnology Research Programme provided support from the early stages of the research. A doctoral scholarship from the New Zealand Accident Compensation Corporation including a stipend and a generous research expenses component was subsequently awarded. Two grants to cover research costs specific to studies 1 and 3 of my project were awarded by the Canterbury

Medical Research Foundation and the Transport Research Education Trust Board respectively. In addition to this I received a number of travel grants to attend three international conferences. These were awarded by The Canterbury Medical Research Foundation, the New Zealand Federation of Graduate Women, Canterbury branch, the Royal Society of New Zealand, Canterbury branch, Universities New Zealand – Te Pōkai Tara, and the Maurice and Phyllis Paykel Trust as well as funds supplied from the Psychology Department of the University of Canterbury.

I would like to thank my family, especially my parents Adrienne and David Hoggarth. My parents have supported me in all my endeavours. I have never felt pressure to act a certain way or achieve a certain goal, but was rather provided with the support required to make my way in whatever pursuit I chose. Finally I would like to thank my partner Alex Mortlock for all his support over these past four years. It has not always been easy, particularly with both of us concurrently undertaking clinical psychology training, but he was always available to bounce ideas off, to give me his opinions on research matters, and to provide some much needed grounding when things became stressful. Thanks.

TABLE OF CONTENTS

Preface

The research for this PhD thesis was carried out between June 2007 and October 2010 while the PhD candidate was enrolled in the Department of Psychology, University of Canterbury. The research was based in the Christchurch Neurotechnology Research Programme, Van der Veer Institute for Parkinson's and Brain Research, and was supervised by Associate Professor Richard Jones, Dr Carrie Innes, and Associate Professor John Dalrymple-Alford.

Studies 1 and 3 (Healthy Older Drivers and Dementia and Driving studies) were conducted in collaboration with the Driving and Vehicle Assessment Service (DAVAS) at Burwood Hospital, Christchurch. Participants for Study 3 were referrals to DAVAS for medical driving assessments due to cognitive impairment, while participants in Study 1 were cognitively-unimpaired older drivers. The off-road experimental assessment of all participants was administered by the PhD candidate throughout the project; blinded on-road driving assessments were conducted by occupational therapists. The experimental assessment of participants included cognitive testing, administration of *SMCTests*, collection of demographic, medical condition, and personality test data and, for the Dementia and Driving study, interviewing of informants.

Aspects of this research were presented by the PhD candidate at the following conferences and meetings: International Neuropsychological Society Annual Meeting (Atlanta, Georgia, February 2009), the 5th International Driving Symposium on Human Factors in Driver Assessment - awarded HONDA Outstanding Student Paper Award (Big Sky, Montana, June 2009), Department of Psychology Postgraduate Research Conference – awarded Best Student Conference Paper (Christchurch, November 2007), Van der Veer Institute Brain Research Forum (Christchurch, July 2008), New Zealand Psychological Society Annual Conference – awarded Best Student Conference Paper Prize (Christchurch, August 2008), University of Canterbury Showcase Postgraduate Conference – awarded Best Presentation Within the College of Science (Christchurch, September 2008), MacDiarmid Young Scientist of the Year Awards (a national science competition for postgraduate students) – awarded runner-up for the *Science and our Society* category (Auckland, July 2009), Chartered Institute of Logistics and Transport (Christchurch, March 2010), Canterbury District Health Board Department of Medical Physics & Bioengineering Seminar (Christchurch, June 2010), University of Canterbury College of Science 'PhD in 3' final (Christchurch, May 2010), Van der Veer Institute Brain Research Symposium (Christchurch, August 2010), and University of Canterbury Showcase Postgraduate Conference (Christchurch, September 2010).

The following publications were generated during this PhD research (two additional papers are expected to be generated on the basis of Chapter 8, and two short papers have been submitted for a conference in 2011):

Full papers

Hoggarth, P.A., Innes, C.R.H., Dalrymple-Alford, J.C., Severinsen, J.E., and Jones, R.D. (2010). Comparison of a linear and a non-linear model for using sensory-motor, cognitive, personality, and demographic data to predict driving ability in healthy older adults. *Accident Analysis and Prevention*, *42*, 1759-1768 (provided in Appendix A)

Hoggarth, P., Innes, C., Dalrymple-Alford, J., Croucher, M., Severinsen, J., Gray, J., Oxley, J., Brook, B., Abernathy, P., and Jones, R. (in press). Assessment of older drivers in New Zealand: the current system, research, and recommendations. *Australasian Journal on Ageing* (uncorrected proof provided in Appendix B)

Short papers

Hoggarth, P., Jones, R., Innes, C., and Dalrymple-Alford, J. (2009). Driving assessment and subsequent driving outcome: a prospective study of safe and unsafe healthy driver groups. *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Big Sky, Montana, USA, 433-439 (provided in Appendix C)

Hoggarth, P., Innes, C., Dalrymple-Alford, J., and Jones, R. (2011). On-road driving assessment errors associated with pass and fail outcomes for older drivers with cognitive impairment. *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Lake Tahoe, California, USA, 351-357 (provided in Appendix D)

Hoggarth, P., Innes, C., Dalrymple-Alford, J., and Jones, R. (2011). Self-rated distress related to medical conditions is associated with future crashes or traffic offences in older drivers. *Proceedings of the 6th International Driving Symposium on Human Factors in Driver*

Assessment, Training, and Vehicle Design, Lake Tahoe, California, USA, 80-86 (provided in Appendix E)

Refereed conference abstracts

Hoggarth, P.A., Innes, C.R., Dalrymple-Alford, J.C., and Jones, R.D. (2009) A nonlinear model of cognitive and sensory-motor test performance can aid in prediction of on-road driving ability in older adults. *Proceedings of the Thirty-Seventh Annual Meeting International Neuropsychological Society*, (11-14 February 2009) Atlanta, GA, U.S.A. p.28 http://medicine.osu.edu/sitetool/sites/pdfs/inspublic/final_program_linked_abstracts.pdf

Hoggarth, P.A., Innes, C.R.H., Jones, R.D., Dalrymple-Alford, J.C. (2008). Cognitive and sensory-motor performance cannot fully predict unsafe driving performance in healthy older drivers. *Proceedings of the New Zealand Psychological Society Annual Conference*, (29-31 August 2008) Christchurch, New Zealand. p. 42

Table of Abbreviations

4IADL	Four-Item Instrumental Activities of Daily Living scale
ADAS-Cog	Alzheimer’s Dementia Assessment Scale – Cognitive Behaviour
ADL-IS	Alzheimer’s dementia Activities of Daily Living International scale
BAI	Beck Anxiety Inventory
CDR	Clinical Dementia Rating
DAS	Driving Anger Scale
DAVAS	Driving and Vehicle Assessment Service, Burwood Hospital, Christchurch
DBQ	Driver Behaviour Questionnaire
DHQ	Driving Habits Questionnaire
DRS-2	Dementia Rating Scale–2
GDS	Geriatric Depression Scale
GP	General Practitioner
IQCODE	The Informant Questionnaire on Cognitive Decline in the Elderly
JOLO	Judgment of Line Orientation
MCI	Mild cognitive impairment
MMSE	Mini-Mental State Exam
MoCA	Montreal Cognitive Assessment
SMCTests™	Sensory-Motor and Cognitive Tests™
SMMSE	Standardized Mini-Mental State Exam
STISIM™	Systems Technology Inc. Simulator
TMT A	Trail Making Test A
TMT B	Trail Making Test B
UFOV®	Useful Field of View®
VR-DR	Virtual Reality Driving Simulator
WTAR	Wechsler Test of Adult Reading

CHAPTER 1 -

Introduction and Statement of the Problem

1.1 Driving and Older Adults

Driving is important to maintain access to one's environment and resources such as supermarkets, shopping malls, banking, and health care. Driving also facilitates social interactions including family and social events. The value of driving does not diminish as we age. In fact, as physical limitations increase, it may become more important. Factors associated with an increased rate of driving cessation within an older adult population include older age (as people age they become more likely to cease driving) (Ragland et al., 2005; Ackerman et al., 2008; Edwards, Bart et al., 2009), poorer physical health (Ragland et al., 2005; Sims et al., 2007), poorer cognitive health (Ackerman et al., 2008; Mezuk & Rebok, 2008; Edwards, Bart et al., 2009), lower levels of education (Ragland et al., 2005; Mezuk & Rebok, 2008), and not being married (Ragland et al., 2005; Mezuk & Rebok, 2008).

Compared to older people who continue to drive, older people who cease driving are more likely to have decreases in physical functioning (Edwards, Lunsman et al., 2009), increases in depressive symptoms (Ragland et al., 2005), decreases in out-of-home and social activity participation (Marottoli et al., 2000; Mezuk & Rebok, 2008; Edwards, Lunsman et al., 2009), increases in the likelihood of entry into long-term care facilities (Freeman et al., 2006), and even increases in mortality (Edwards, Perkins et al., 2009). Older people are more at risk than middle-aged people of being killed and injured, particularly as pedestrians (Evans, 2000; Ministry of Transport, 2008a). For many older people accessing alternative transportation may not be possible due to their location, expense, or mobility problems, or unavailability of friends or family members to provide transportation.

Given the increase in negative outcomes for older drivers who cease driving and the difficulties associated with attaining acceptable alternative transportation, it is of the utmost importance that older people continue to drive as long as they can. From a public policy perspective cessation should only be required if there is evidence to believe that the person is putting themselves and/or other road users at seriously increased risk.

1.2 The Need for Prediction of On-Road Driving Ability

On-road driving assessment is a widely accepted 'gold standard' in determining on-road driving ability. However, there are inherent risks to driving assessors and the public in allowing people with physical or cognitive problems to be assessed on public roads. On-road driving assessment is a limited and expensive resource. In most areas of New Zealand, on-road assessments are performed at the private cost of the individual. With the expected increase in older adults in the coming decades (Organisation for Economic Co-operation and Development, 2001), the issue of detecting potentially unsafe older drivers has become a pressing issue. Accurate prediction of driving ability using off-road testing will lead to more efficient use of limited assessment resources and prevent unnecessary evaluation of drivers whose risk of an adverse driving event is low.

In New Zealand, mandatory on-road driving tests for drivers aged 80 and over was abolished in December 2006 and were replaced by compulsory licence renewal ages of 75, 80, and every two years thereafter. This change was due in large part to pressure by older adult organizations who claimed the policy was ageist (see Sullivan, 2004 for a response to Grey Power's criticism of Land Transport New Zealand's statistics that showed older adults were at increased risk of being injured and killed on the road). The abolition of compulsory testing raised inquiry as to whether at-risk older drivers could be identified using off-road testing. An accurate screening test for potential driving problems would be particularly valuable at licence renewals.

In a related context, a New Zealand study determined a classification model to detect those people with brain disorders who were likely to receive a fail score on an on-road driving assessment (Innes et al., 2007). This model was constructed using 50 participants and is currently used in several occupational therapy practices in New Zealand as part of the formal driving assessment process. However, a follow-up study of 200 drivers with brain disorders or suspected dementia found that classification of Pass and Fail was more accurate when the dementia group was considered separately from the brain disordered group, rather than assessing all participants as one sample (Innes, Jones, Dalrymple-Alford et al., 2009). This suggests that a selection of measures could be optimized for drivers with suspected dementia to provide better prediction of driving ability compared to the model constructed using a brain disordered sample.

1.3 Objectives

This research project addressed the following objectives:

1. Review the effects of normal ageing on driving performance.
2. Review how mild cognitive impairment and dementia affect driving performance.
3. Review how driving ability is currently assessed.
4. Assess the accuracy of measures currently used to assess driving ability.
5. Determine measures that predict on-road driving performance in a group of older cognitively-unimpaired drivers.
6. Follow cognitively-unimpaired older drivers to determine if drivers who fail an on-road driving assessment are more likely to experience adverse driving events than drivers who pass the assessment.
7. Determine measures the predict on-road driving performance in a group of drivers with mild cognitive impairment and dementia.

CHAPTER 2 - Effects of Ageing, Mild Cognitive Impairment, and Dementia on Driving Performance

2.1 Objectives

The objectives of this chapter are to investigate the effects of normal ageing, mild cognitive impairment and dementia on driving performance. Before investigating how mild cognitive impairment, and dementia affect driving, it is first necessary to define how they are diagnosed. Following this, the current driver licensing requirements in New Zealand will be reviewed, particularly for older drivers.

2.2 Normal Ageing

There are a number of changes associated with ageing that can be grouped broadly into visual, cognitive, and physical domains (Janke, 1994; Staplin et al., 1998; Anstey et al., 2005). Visual factors include physical changes in the sensory mechanisms of the eye, such as the development of cataracts, which interfere with functional mechanisms and lead to reduced visual acuity, reduced contrast sensitivity, visual field loss, deficits in depth and motion perception, increased glare sensitivity, and poorer vision in low light. Physical factors include decreased muscle strength and endurance, slower reaction and movement times of limbs, and reduced physical mobility, particularly in the trunk, neck and head. Cognitive factors include decrements in processing speed, memory function, planning, and decision-making, visual scanning, focused attention on environmental stimuli, and dividing attention between multiple tasks.

With ageing also comes an increase in the incidence of chronic disease of which vascular risk factors have been associated with impaired cognitive processes (Raz et al., 2007).

2.3 Dementia and Mild Cognitive Impairment

2.3.1 Dementia

Dementias are characterized by the development of multiple cognitive deficits which generally include memory deficits. The common dementias are progressively degenerative in nature but all must be associated with significant impairment in social and/or occupational functioning in order to meet diagnostic criteria. Invariably, personality and behaviour become altered to varying degrees. There are two commonly used diagnostic criteria for the diagnosis of dementia. A general one is found in the cognitive disorders chapter of the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV-TR) (American Psychiatric Association, 2000). The second was proposed by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) (McKhann et al., 1984) and is specific to the diagnosis of Alzheimer's dementia.

Dementias in the DSM-IV-TR are subclassified into diagnoses of dementia of the Alzheimer's type, vascular dementia, dementia due to other general medical conditions (e.g., HIV infection, head injury, Huntington's disease, Lewy-body dementia, frontotemporal dementia, multiple sclerosis), substance-induced persisting dementia, dementia due to multiple aetiologies, and dementia not otherwise specified. All dementias in the DSM-IV-TR share a common symptom list with individual subtypes differentiated on the basis of their aetiology. Diagnostic criteria for dementia of the Alzheimer's type are reproduced below:

A. The development of multiple cognitive deficits manifested by both:

1. Memory impairment (impaired ability to learn new information or to recall previously learned information)
2. One (or more) of the following cognitive disturbances:
   - (a) aphasia (language disturbance)
   - (b) apraxia (impaired ability to carry out motor activities despite intact motor function)
   - (c) agnosia (failure to recognize or identify objects despite intact sensory function)
   - (d) disturbance in executive functioning (i.e., planning, organizing, sequencing, abstracting)

B. The cognitive deficits in criteria A1 and A2 each cause significant impairment in social or occupational functioning and represent a significant decline from a previous level of functioning.

C. The course is characterized by gradual onset and continuing cognitive decline.

D. The cognitive deficits in Criteria A1 and A2 are not due to any of the following:

(1) other central nervous system conditions that cause progressive deficits in memory and cognition (e.g., cerebrovascular disease, Parkinson's disease, Huntington's disease, subdural hematoma, normal-pressure hydrocephalus, brain tumor)

(2) systemic conditions that are known to cause dementia (e.g., hypothyroidism, vitamin B or folic acid deficiency, niacin deficiency, hypercalcemia, neurosyphilis, HIV infection)

(3) substance-induced conditions

E. The deficits do not occur exclusively during the course of a delirium.

Alzheimer’s dementia is the most common form of dementia, accounting for half to two-thirds of all cases (Lau & Brodney, 2008). Surveys of all dementias have reported prevalence rates of approximately 1% for ages 60 to 70, 7% between ages 70 and 79, 14% for ages 80 to 84, 43% for ages 85 and 89, and 31% to 65% for those aged 90 and above (Ritchie & Kildea, 1995; De Ronchi et al., 2005; Plassman et al., 2007). The rate and sequence of progression is unpredictable at an individual level, although early stages of degeneration in Alzheimer’s occur in cortical and hippocampal medial temporal lobe areas with cell loss generally spreading rostrally to prefrontal and parietal areas (Lezak et al., 2001). Alzheimer’s dementia is a diagnosis that can be given once there is supporting evidence (history or early memory decline, formal cognitive testing, and/or neuroimaging) and when alternative aetiologies are deemed unlikely. One important differential diagnosis is primary vascular dementia which is due to cerebrovascular insult most commonly caused by repeated ischaemic strokes. Medical conditions such as Multiple Sclerosis, dementia with Lewy bodies and Frontotemporal dementia must also be ruled out, as well as other less common dementias (e.g., substance abuse, or caused by HIV infection). Additional specifiers available in the diagnosis of Alzheimer’s in the DSM-IV-TR include whether behavioural disturbance is present (e.g., wandering, agitation) and whether the disease has early or late onset (early onset is for those aged 65 years or below when symptoms became apparent).

The NINCDS-ADRDA criteria are widely used in research for the diagnosis of Alzheimer's dementia. These criteria differ from the DSM-IV-TR in that they have three levels with which Alzheimer's can be diagnosed. The 'Possible' level requires that criteria for Alzheimer's are met, but with the presence of other possible influences. The 'Probable' level requires that criteria for Alzheimer's are met with no other apparent influences for impairment. The 'Definite' level requires that histopathologic evidence of Alzheimer's is obtained from biopsy or autopsy. Simplified criteria for 'Probable' Alzheimer's dementia are reproduced below (McKhann et al., 1984):

1. Dementia established by clinical examination and confirmed by neuropsychological tests.
2. Progressive worsening of memory.
3. Deficits in two or more areas of cognition.
4. No disturbance of consciousness (i.e., not present solely during the course of delirium)
5. Onset between ages 40 and 90, most often after age 65.
6. Absence of systemic disorders or brain diseases that, in and of themselves, could account for the progressive deficits in memory and cognition.

Additional supporting evidence for a diagnosis of Alzheimer's dementia includes impaired activities of daily living, family history of similar disorders, and normal results for a lumbar puncture and EEG, or evidence of cerebral atrophy using CT imaging with progression noted on serial measurements. While activities of daily living deficits are not mentioned in the specific criteria, general consensus (including DSM-IV-TR criteria) is that activities of daily living must be impaired for a diagnosis of Alzheimer's dementia to be given, with an absence of deficits in this area more likely to lead to a diagnosis of mild cognitive impairment (see Section 2.3.2).

Neither the DSM-IV-TR nor the NINCDS-ADRDA criteria provide guidelines for determining the severity of dementia. Severity is often established in research using the Clinical Dementia Rating, with a score of 0 indicating no dementia, a score of 0.5 indicating very mild impairment (sometimes used to denote mild cognitive impairment, which is discussed in Section 2.3.2), a score of 1 indicating mild impairment, a score of 2 indicating moderate impairment, and a score of 3 indicating severe impairment. In the absence of the CDR or other alternatives such as the Dementia Rating Scale (Jurica et al., 2001), the Mini-

Mental State Exam can provide rough ranges to indicate severity, with a score of ≤9 indicating severe impairment, 10-20 indicating moderate impairment, 21-24 indicating mild impairment, and 25 and above falling within the ‘intact’ range (Mungas, 1991). Further discussion of the use of MMSE for detecting dementia is in Section 5.7.1.

In research studies, the diagnosis of dementia is generally made following extensive cognitive testing to document the degree of cognitive impairment across the specific cognitive domains mentioned in the DSM-IV-TR or NINCDS-ADRDA criteria. In addition, enquiries are made of spouses or family members to confirm that current everyday functioning is a decline from previous levels of functioning. Cognitive measures used in the diagnosis of dementia in the Dementia and Driving study are discussed in Section 8.1.3.

2.3.2 Mild Cognitive Impairment

Mild cognitive impairment (MCI) is proposed as a stage of cognitive impairment that lies between the normal cognitive changes expected in ageing and the pathological changes seen in dementia (Petersen et al., 2001; Petersen, 2004). MCI has increasingly gained acceptance as a precursor state to dementia. Carefully controlled studies suggest that people with a diagnosis of MCI are more likely to develop dementia than same-aged non-cognitively impaired people, with an average rate of progression of MCI to Alzheimer’s dementia of 14% per year (Petersen et al., 2001). A diagnosis of MCI made using more recent criteria (Petersen, 2004) requires that a minimum of one area of cognition is affected, which does not have to be memory, with activities of daily living not significantly affected. Petersen (2004) provided a flowchart (Figure 2-1) that separates the diagnosis of MCI into four categories: Amnestic MCI single domain where there is only significant decline in memory processes, and Amnestic MCI multiple domain where there is also significant decline in at least one other area of cognition. There is also a Non-Amnestic MCI classification for people who lack memory impairment but have one (single domain) or more than one (multiple domains) areas of deficit in other cognitive processes.

Although few researchers now agree, Petersen (2004) suggested that specific subtypes of MCI may be more likely to develop into specific subtypes of dementia, with amnestic types more likely to become Alzheimer’s dementia and types characterized by non-memory impairment more likely to become dementia types such as Dementia with Lewy Bodies or Frontotemporal Dementia.

As there is no universally accepted set of diagnostic criteria for MCI, the most widely used criteria for Amnestic MCI (Petersen, 2004) are, in simplified form:

1. Subjective memory complaint usually corroborated by an informant.
2. Objective memory impairment for age.
3. Essentially preserved general cognitive function.
4. Largely intact functional activities.
5. Not demented.

The criterion regarding functional activities leaves ample room for subjective decisions regarding whether the level of impairment is indicative of MCI or dementia.

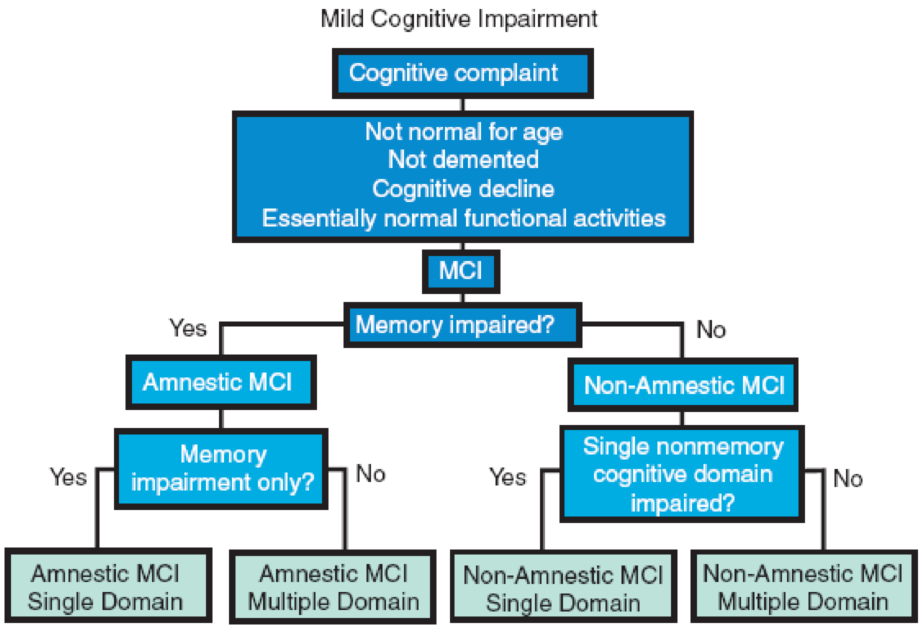


Figure 2-1. Proposed flowchart showing the four different types of MCI and how the diagnostic decision is arrived upon. Adapted from Petersen (2004).

Despite the diagnosis of MCI being somewhat variable across researchers at this point, the detection of early stages of cognitive impairment is very important as the affected person and their family need to make plans for their future and it may determine the administration of any current or future treatment that may be most effective in the early stages of a possible dementia.

2.4 Driving with Ageing, Dementia, and Mild Cognitive Impairment

2.4.1 Normal Ageing and Driving

The visual, cognitive, and physical factors listed in Section 2.2 were reviewed due to their possible negative affects on driving (Janke, 1994; Staplin et al., 1998; Anstey et al., 2005). Specific medical conditions have been linked to negative driving outcomes including heart disease (McGwin et al., 2000; Sagberg, 2006), high blood pressure (Anstey et al., 2005), stroke (McGwin et al., 2000; Sims et al., 2000; Sagberg, 2006), and dementia (Cooper et al., 1993; Zuin et al., 2002). Marottoli et al. (1994) found that older drivers with a higher number of chronic conditions were more likely to be involved in 12-month prospective self-reported crashes. Molnar et al. (2007) found that being "bothered a great deal by diabetes mellitus" was associated with increased motor vehicle crashes in older drivers. The authors suggest this demonstrates that functional effects of an illness may be a more useful correlate of crash incidence than the simple presence of an illness. That is, the severity and control of a medical condition may tell us more than simply whether the person has the disease.

Older adults with intact cognition may be able to compensate for age-related changes provided they occur gradually. But this may not be the case. Lövsund, Hedin, and Törnros (1991) found that only 4 of 31 participants with visual field deficits were able to perform as well as a control group in detecting visual stimuli during a simulated drive. It is unclear how an older driver could compensate for the effects of illness such as heart disease or high blood pressure which presumably affect driving through wear and tear on the vascular system and have been linked with impaired cognitive processes (Raz et al., 2007).

The evidence for an effect of age-related changes to the visual system on driving safety is not strong. Anstey et al. (2005) performed a review of the research literature from 1991 to 2002 and concluded that the effect of reduced visual acuity on the outcome of on-road assessment or real-world crashes is inconsistent between studies. They reported weak associations between contrast sensitivity and on-road outcomes, and no evidence for the association between peripheral visual fields or depth perception and on-road outcomes. Anstey et al. (2005) suggest that visual tests used in isolation are not strong predictors as they do not tap into the visual and cognitive complexity of the driving task.

2.4.2 Driving with Alzheimer's Dementia

The diagnosis of dementia is not a sensitive or specific enough measure on its own to determine driving ability. Many people with early dementia are able to pass an on-road driving assessment with pass rates ranging from 37% to 73% (Fox et al., 1997; Grace et al., 2005; Lincoln et al., 2006). The task of deciding which drivers with dementia are safe to continue driving is therefore a difficult matter. A review of several published guidelines for assisting in the decision-making process regarding which drivers with dementia are safe and which are unsafe reveals several similarities. Recommendations that people with moderate and severe dementia cease driving have been made (Canadian Medical Association, 2006; Australian and New Zealand Society for Geriatric Medicine, 2010) with an accompanying suggestion that people with mild dementia may be able to continue driving with appropriate monitoring and assessment. Statements that results of neuropsychological tests cannot be used reliably thus far to determine which drivers with dementia are safe and unsafe on the road have also been made (Australian and New Zealand Society for Geriatric Medicine, 2010; Iverson et al., 2010). The American Academy of Neurology identified the Clinical Dementia Rating (CDR) as the most useful measure of overall cognitive decline (Iverson et al., 2010). They provide further recommendations to take into consideration caregiver ratings of poor driving, the incidence of traffic offences and crashes, and to look for changes in driving patterns such as reduced mileage and situational avoidance as potential risk factors. Other literature reviews and research studies support the use of the CDR as a global measure of dementia severity that can be useful in deciding whether a patient is likely to have problems with driving (Johansson et al., 1996; Hunt et al., 1997; Duchek et al., 1998; Dubinsky et al., 2000; Brown & Ott, 2004; Ott et al., 2008).

The New Zealand Transport Agency guide for New Zealand medical practitioners (New Zealand Transport Agency, 2009) provides little specific guidance, with no endorsed tests to detect dementia, no mention of degrees of dementia severity and how they relate to driving safety, and an extreme and ill-informed statement that "Individuals with confirmed dementia or cognitive impairment from whatever cause should not drive" (p. 33). This statement clearly goes against evidence that many people with diagnosed dementia can pass an on-road driving assessment, and is also contrary to the advice of the other professional bodies, listed above, which provide a more detailed and objective analysis to testing driving ability in those with dementia.

2.4.3 Driving with Mild Cognitive Impairment

Due to the lesser severity of MCI compared to Alzheimer's dementia, it would be expected that MCI would have less effect on driving ability. There are few studies that have investigated driving ability in samples of people with MCI. Frittelli et al. (2009) tested a group of 20 participants with Alzheimer's dementia, 20 with MCI, and 19 controls on a driving simulator. The Alzheimer's participants received poorer scores than the control group on the time it took to complete the drive, time to collisions, number of off-road events, and visual reaction times. The MCI group were poorer than controls only in mean time to collision. Other studies have not classified their participants as MCI but rather as "very mild dementia" with a Clinical Dementia Rating of 0.5. Those drivers with a CDR score of 0.5 performed more poorly on an on-road test than a non-impaired control group (Brown et al., 2005). Duchek et al. (1998) found that on-road test scores dropped as dementia severity increased from CDR 0 (unimpaired), to 0.5, to 1. Survival curves for failing repeated on-road tests over time were steeper for CDR 0.5 than CDR 0, but less steep than for people with a CDR rating of 1, although these differences were not statistically significant (Duchek et al., 2003). O'Connor et al. (2010) followed a group of cognitive healthy older adults and a group with cognitive impairments suggestive of MCI over a period of five years with self-ratings of a number of driving behaviours. They found a significant increase in ratings of driving difficulty over time in the non-amnestic and multiple domain MCI types compared to the control and amnestic MCI groups. The authors suggest that functional loss of complex behaviours such as driving occurs in those with MCI. This suggests that the early detection of cognitive impairment in older drivers is an important clinical objective.

More studies are needed that utilize MCI samples in order to determine the level of risk in this group compared to those with a dementia diagnosis.

2.4.4 Summary

Normal ageing heralds many changes in visual, cognitive, and physical domains as well as an increase in the incidence of chronic illness, comorbidity, and cognitive disorders. While older drivers without cognitive impairment may be able to adjust to some extent to these gradual changes over time, drivers with cognitive impairment caused by MCI or Alzheimer's dementia are at a disadvantage. The task of predicting driving ability is made more difficult by the idiosyncratic nature of MCI and Alzheimer's dementia, with some people able to continue driving safely with others who are seemingly similarly impaired having lost this ability. The lack of research into the effect of MCI on driving ability is a concern as we can expect a large increase in the numbers of older drivers with MCI in the coming decades.

There is agreement that drivers with dementia will at some point become unsafe drivers. However, a reliable way to measure when this occurs is not readily forthcoming. Finding measures that help to separate those drivers with MCI and Alzheimer's dementia who can drive safely from those who cannot is necessary for developing ways of screening for driving ability in a primary care setting.

2.5 Crash and Traffic Offence Risk

2.5.1 General Public

Crashes are a rare event, particularly crashes in which people are injured or killed. In New Zealand, the Ministry of Transport keeps record of crashes that have resulted in injury or death within 30 days. This crash database includes any crash that occurred on a public road which was attributable to a motor vehicle or its load. In 2009, 384 people died on New Zealand roads as a result of a motor vehicle crash and 14,541 were injured (Ministry of Transport, 2010). This equates to death rates of 1.2 per 10,000 registered vehicles, and injury rates of 45 per year per 10,000 registered vehicles.

The majority of crashes do not result in an injury or death and, hence, are not found in official records. A New Zealand study of 853 drivers aged 15 to 84 (mean 39.3 years) found a rate of self-reported crashes of 38.2% over 5 years, or 7.6% a year (Sullman & Baas, 2004). An Australian study found a 12-month self-reported crash rate of 10.8% for a group of 443 employees of an insurance company with an age range of 18 to 68 years (mean age 44 years) (Davey et al., 2007). It is clear that self-reported crash rates are higher than the incidence of police-recorded crashes that result in death and injury. Although most self-reported crashes would be considered less serious than those officially-recorded, they could none-the-less tell us about the on-road safety of drivers.

Traffic offences are another measure that can tell us about driving safety and have the advantage of having higher base rates than officially-recorded crashes. 17.6% of an Australian sample of 443 drivers self-reported receiving fines and demerit points due to traffic offences in a 12-month period (Davey et al., 2007). For the 12 months ended 30 June 2008, New Zealand police issued 760,720 speeding offences to a licensed population of just over 3 million drivers (New Zealand Police, 2008; Ministry of Transport, 2009). This represents a speeding offence for 24% of drivers, although this is an over-estimation as the data do not account for single drivers receiving multiple offences.

The commission of traffic offences has been associated with an increased likelihood of having crashes in general population samples (Rajalin, 1994; Parker et al., 1995; Cooper, 1997). People observed driving at higher speeds on public roads are more likely to have a state-recorded history of crashes and traffic offences over the previous seven years (Wasielewski, 1984). The higher base rate of traffic offences compared to crashes, and the relationship between traffic offence and crash risk indicates that traffic offences could be a useful measure for detecting drivers with unsafe driving behaviours.

2.5.2 Older Drivers

Interpretation of older driver involvement in death and injury causing crashes is confounded by several factors. The most influential factor is that older drivers are more physically fragile than younger drivers and hence more likely to be killed or injured than a younger driver in a crash of equivalent severity (Organisation for Economic Co-operation and Development, 2001; Li et al., 2003; Meuleners et al., 2006; Tefft, 2008). Li, Braver, and Chen (2003) developed a method for statistically controlling for the influence of physical fragility for death statistics and found substantially decreased death rates following data correction. Statistics for New Zealand driver deaths, both uncorrected and corrected for fragility, are shown in Figure 2-2. The figure clearly shows that controlling for fragility reduces driver deaths in those aged 60 and over.

Another confounder of older adult crash rates is the custom of representing crash data per km driven, as in Figure 2-2, which presents risk as if all drivers were equally exposed, i.e., all drove the same number of km. Older drivers, on average, drive fewer km per year than middle aged drivers (Evans, 2000). Two population based studies of Finnish and Dutch drivers found that driving fewer km was associated with increased accident involvement across all age groups (Hakamies-Blomqvist et al., 2002; Langford et al., 2006). But the relationship between km driven and crashes appears to be even more complicated than this. Drivers aged 75 and over who reported driving fewer than 3,000 km a year have been shown to have a positive age to crash risk correlation, while drivers of the same age who drove more than 3,000 km a year had the lowest crash rates for all age groups (Langford et al., 2006).

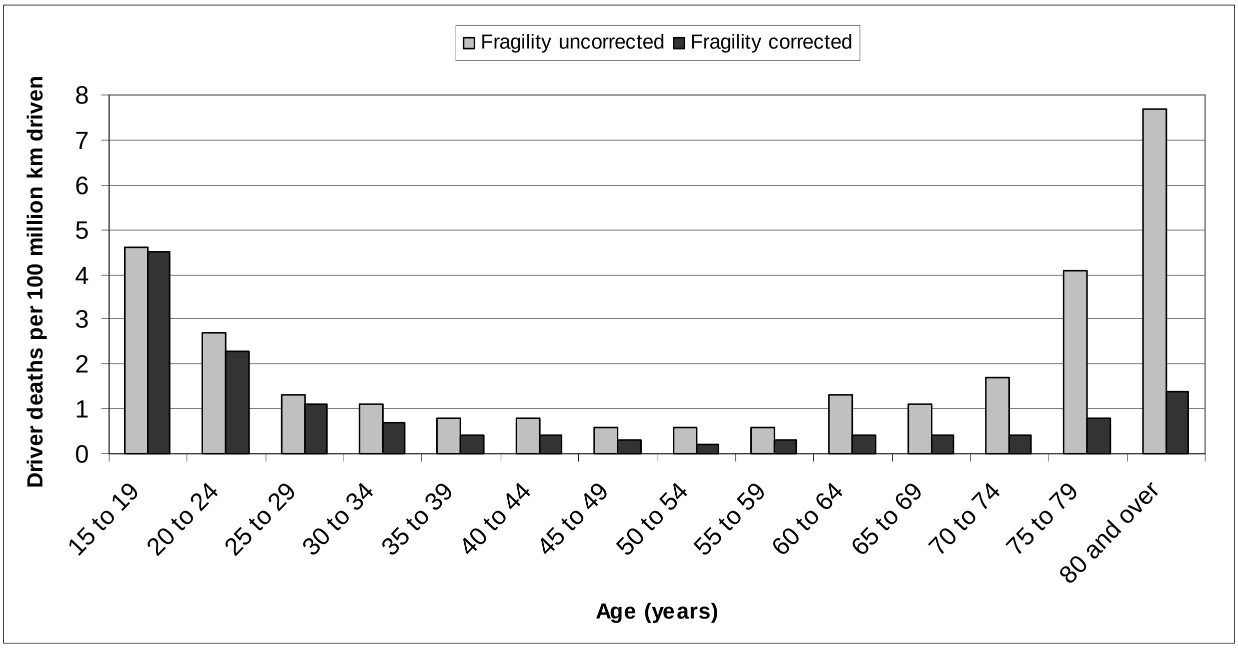


Figure 2-2. Deaths rates of drivers involved in crashes both uncorrected and corrected for fragility. Adapted from Ministry of Transport (Ministry of Transport, 2008a)

The authors suggest this could be due to an over-representation of physically and cognitively-impaired drivers in the low-mileage oldest age group. This pattern indicates that reducing driving frequency and distance could act as a risk factor for negative driving outcomes. There is likely also an affect for km driven in that those who drive more frequently are on the road more often and thus exposed to the risk of a crash more than those who drive less frequently.

Presenting death or injury statistics by the number of licensed drivers in each age group rather than by per km travelled essentially controls for differing driving exposure between age groups (Figure 2-3). The shape of the distribution is similar to the distribution of deaths per km driven but with a shallower rise in the oldest age groups. Unfortunately the data in Figure 2-3 is not adjusted for fragility but it would be expected to follow a similar reduction for older drivers as shown in Figure 2-2 due to the outcome measure being injuries and deaths which are mostly influenced by fragility.

Evans (2000) investigated the threat posed to other road users by relicensing male drivers of different ages. He found that relicensing an 80 year old male driver represented 26% less threat to other road users than relicensing a 40 year old male driver. Relicensing a 20 year old male represented a threat of 140% greater than relicensing an 80 year old male(Evans, 2000).

The lower rates of threat associated with older drivers were due to their reduced driving exposure compared to the younger driver groups (Evans(2000).

The number of drivers injured and killed on New Zealand roads in 2008 as a percentage of licensed drivers in each age group over 60 years old were 0.24% for ages 60-64, 0.23% for aged 65-69, 0.25% for aged 70-74, 0.30% for ages 75-79, and 0.45% for ages 80 and above (Ministry of Transport, 2009). A study of New Zealand drivers aged 80 and over found crashes resulting in injury and death in 0.8% of the sample over two years, which is consistent with official crash rates reported previously (Keall & Frith, 2004a). Self-reported crash rates, which include those not severe enough to have been recorded officially, range from 4% to 10% per year in older drivers (Marottoli et al., 1994; Sullman & Baas, 2004; Anstey et al., 2009).

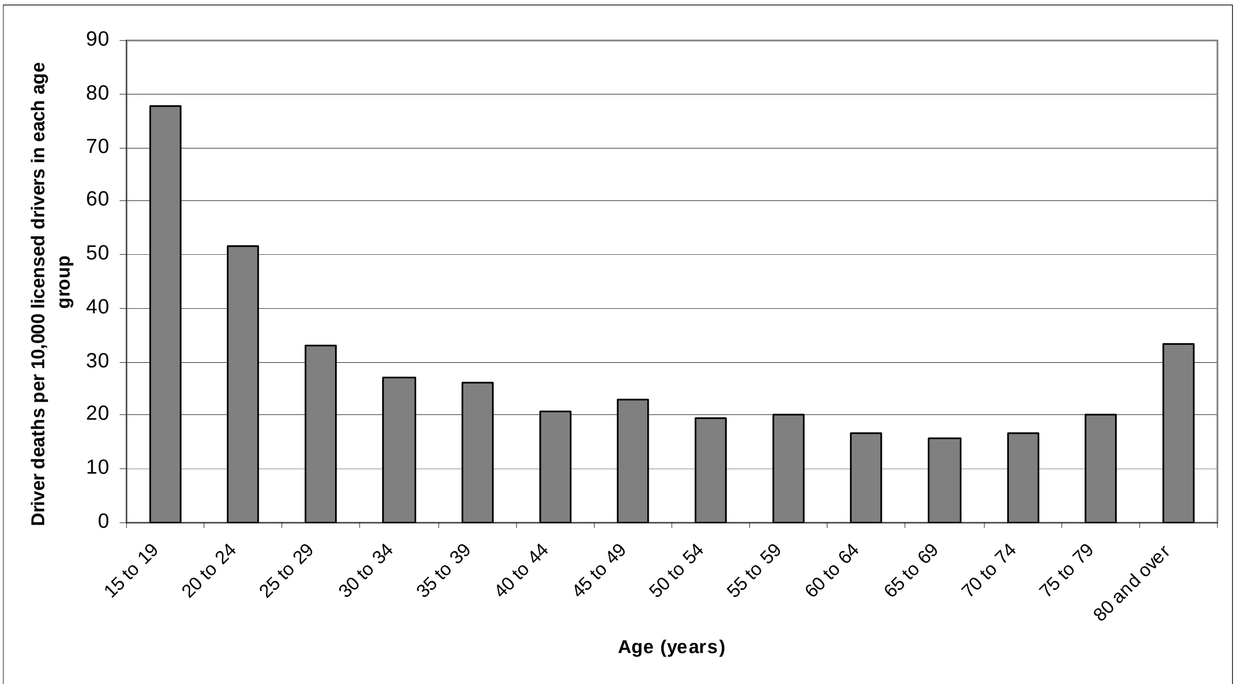


Figure 2-3. Deaths rates of drivers involved in crashes per 10,000 drivers licences in each age range. Adapted from a Ministry of Transport report (Ministry of Transport, 2008b)

There is no published official data relating the age-breakdown of traffic offences in New Zealand and few studies have gathered information on self-reported offences. Two studies that collected self-reported data on traffic offences for older drivers found offence rates per year of 2.6% and 3.0% respectively (Hoffman & McDowd, 2010; Oxley et al., 2010). This rate is lower than the self-reported crash rate reported above and may not be a satisfactory estimate of actual traffic offence rates in older drivers.

It is easy to see how the effects of fragility and low-mileage bias lead to beliefs that older drivers are inherently more dangerous on the road. In general, older drivers pose little risk to other road users but are at greatly increased risk of being killed or injured themselves if they are involved in a crash, even if they did not cause it. Nevertheless, regardless of whether injury and death rates are displayed by km driven as in Figure 2-2 or by number of licensed drivers as per Figure 2-3, drivers in the oldest age groups show an increasing trend for death and injury even after corrections for physical fragility are applied. Li et al. (2003) and Meuleners et al. (2006) independently found a proportionate reduction in the effects of fragility in drivers aged 80 and over with a concomitant rise in death and injury rates attributable to increasing frequency of crash involvement. This suggests that the oldest drivers may begin to have crashes more frequently. An increased crash rate is likely related to age-related changes that include the effects of normal ageing, medical conditions, and cognitive disorders.

Meuser et al. (2009) found support for the effect of medical conditions on driving when comparing 4,100 drivers (mean age 80 years) who had reported medical conditions that may affect driving compared to 11,615 age and sex-matched controls. Official records of crashes that involved death, injury, or property damage of US$500 and above were collected for a 14 year period, with 48.7% of the medically-impaired group having a crash in this period versus 27.0% of the control group. Furthermore, 20.7% of the medically-impaired group had two or more crashes in this time versus 7.3% of the control group. The medically-impaired group consisted of people with dementia or other cognitive disorders (45%), vision problems (31%), musculoskeletal/neuromuscular problems (28%), disorders of consciousness (not defined, 16%), cardiac/cardiovascular problems (12%), brain injury, tumour, or stroke (10%), psychiatric problems (8%), and alcohol and drug abuse (3%). This study supports the assertion that it is older drivers with medical conditions who are likely skewing the crash statistics of older drivers.

2.5.3 Alzheimer's Dementia and Mild Cognitive Impairment

The biases that apply to older drivers and official injury and death rates also apply to drivers with dementia and MCI as the majority of those affected are aged 60 and above. The best way to control for this effect is to have a same-aged control group with which to compare the frequency of crashes. Cooper, Tallman, Tuokko, and Beattie (1993) found that drivers diagnosed with dementia had almost 2.5 times as many crashes that resulted in insurance claims versus an age-matched community control group. Zuin, Ortiz, Boromei, and Lopez (2002) found that drivers with dementia (mostly Alzheimer's) were 10.7 times more likely to have a crash than an age-matched control group. Finally, LaFont et al. (2008) followed just under 1000 drivers with a mean age of 72.8 years, some of whom were diagnosed with dementia at baseline and some who were diagnosed two years later. Those who were diagnosed with dementia at the two-year follow-up (but not at the baseline assessment) were 3.4 times more likely to have self-reported a crash in the 5 years *prior* to the baseline assessment than other drivers without central nervous system disease (defined as a diagnosis of Parkinson's disease, head trauma, or stroke). This result suggests that crash rates may begin to rise years before criteria for a clinical diagnosis of dementia are met, which adds to the need to examine MCI in addition to mild dementia.

Drivers with dementia may have their crash risk further raised by adverse side-effects from prescription medications. Rapoport et al. (2008) found that drivers with dementia who were prescribed antipsychotics, antidepressants, or benzodiazepines were 1.5 times more likely to be involved in a crash than a control group of drivers with dementia who were not taking these medications. Another study, however, compared the police-reported crash and violation rates of 143 drivers with Alzheimer's to 715 age-matched controls and found no increased rate of crashes in the Alzheimer's group for the period 1986 to 1993 (Trobe et al., 1996). The authors suggest that this could be due to reduced mileage by the Alzheimer's group which would have limited their exposure to risky driving situations. The higher crash rates found by many studies have led several research groups to suggest that drivers with dementia be assessed on-road every 6 months (Fox et al., 1997; Dubinsky et al., 2000; Duchek et al., 2003; Adler et al., 2005; Frittelli et al., 2009).

No studies were found that compared the crash rates of people with MCI to a control or dementia group. It would be expected that crash rates for this group would fall between those for controls and dementia groups as do their scores on cognitive tests. Likewise no studies were found that reported rates of traffic offences for either dementia or MCI groups. When studies do investigate driving outcomes in these clinical groups, the dependent variables most commonly used are on-road driving performance, simulated driving performance, and crash rates.

2.5.4 Summary

When the effects of physical fragility and low-mileage bias are controlled for, older drivers can be seen to be the safest drivers on the road. However, older drivers are at a higher risk of serious injury and death if involved in a crash due to their greater physical fragility. It is clear that a sub-group of older drivers are more at risk of being involved in crashes due to medical and cognitive disorders (e.g., dementias, delirium, and amnestic disorders). This group should be the prime target of initiatives that seek to detect older drivers who pose undue risk to themselves and others.

2.6 Driving Assessment in New Zealand

2.6.1 General Population

All New Zealand drivers undergo a compulsory screening assessment for visual acuity and peripheral vision at 10-year licence renewals. Drivers aged 75, 80, and every two years thereafter must obtain a medical-fitness-to-drive certificate from their primary care physician in order to renew their licence. This process is explored in detail in Section 2.6.3.

2.6.2 Medical Conditions

A guide for medical practitioners, optometrists, and occupational therapists is produced by the New Zealand Transport Agency to aid in assessment of medical fitness to drive (New Zealand Transport Agency, 2009). Drivers of any age are required to be medically fit to drive, with a range of medical conditions serving as potential contra-indicators of safe driving for personal and commercial drivers. There are explicit visual standards that all drivers must meet in order to obtain a licence but most other medical conditions do not have specifically defined criteria for licensing. It is the responsibility of the medical practitioner to make a decision regarding driving safety or to refer their patient to a specialist who can advise on suitability for driving.

Assessments for medical fitness to drive can be requested for drivers of any age and are often requested for people who have had strokes or brain injuries to determine if and when they may be able to return to driving. An examination of the ages of drivers reported for medical conditions following the introduction of voluntary reporting laws in the US state of Missouri in 1999 showed that 93% of reports were for drivers aged 50 and over, with 83% for drivers aged 70 and over (Meuser et al., 2009). In a sample of 501 drivers referred for medical driving assessments due to brain problems (traumatic brain injury, neurological conditions including stroke and dementia) in three driving assessment practices in New Zealand, 64% were aged 65 years and over (C. Innes, personal communication, January 2011). It may be safe to assume that the majority of drivers referred for medical driving assessments are in the older age groups.

Apart from general information about the effects of different medical conditions on driving presented in the Medical Aspects of Fitness to Drive handbook (New Zealand Transport Agency, 2009), there are no references or statistics supplied to indicate the effect of particular medical conditions on driving ability. It is therefore the decision of the primary care medical practitioner to decide whether their patient is likely to be a safe driver or to refer on for additional assessment.

2.6.3 Older Drivers

The Medical Aspects of Fitness to Drive handbook contains a specific section for medical conditions more likely to affect older drivers. These conditions include early onset of fatigue, slowed responses, visual problems, impaired cognitive function, impaired mobility, and dementia (New Zealand Transport Agency, 2009, p 105). A medical practitioner is asked to take special note of medical history, mental function, vision, cardiovascular and central nervous systems, and the locomotor system in older drivers seeking to renew their licence. The NZTA offers a short test of road sign recognition in the Medical Aspects of Fitness to Drive manual to use with people who may be suspected of having dementia (New Zealand Transport Agency, 2009, p 135). The instructions read “If you suspect a person may be showing signs of forgetfulness or memory loss, give them this simple test on common traffic signs. A person who has trouble with this test or takes a long time to answer may need further assessment.” Unfortunately this test has not been investigated for its relationship to driving outcomes and no score is provided to determine whether the person may have a problem with driving.

There is a specific process for renewing an older driver’s licence, depicted in Figure 2-4.

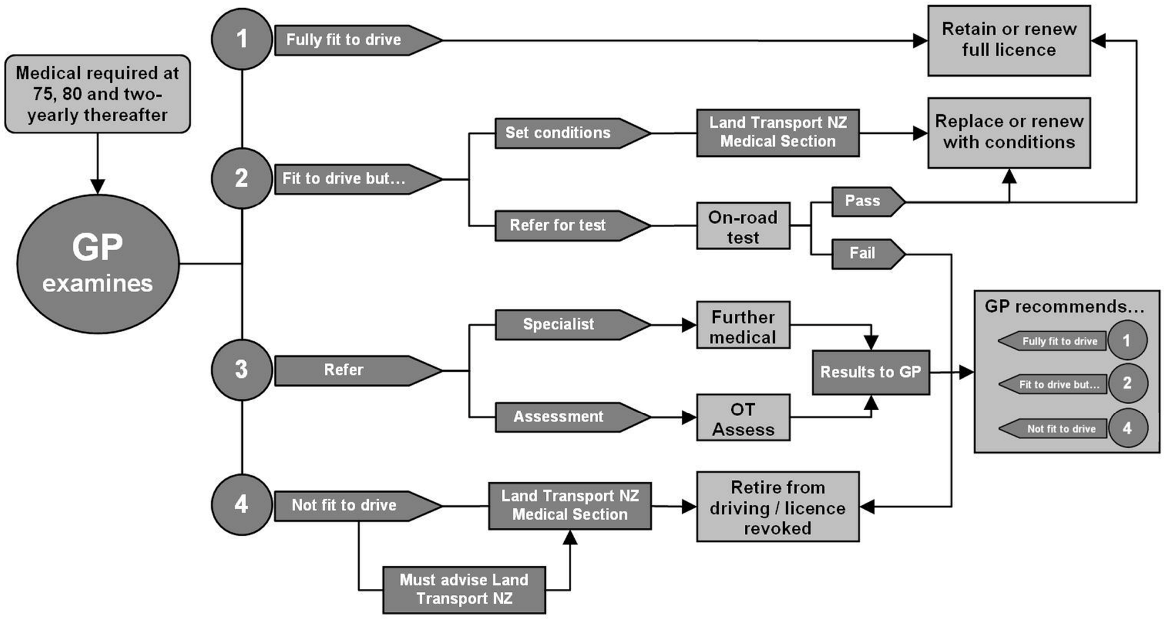


Figure 2-4. Flowchart of the older driver licence renewal system (Adapted from NZTA information pack for general practitioners of December 2006).

The chart depicts four decision pathways. The most direct options (1 and 4) are for a medical practitioner to independently decide whether a patient is medically fit to drive or not and to recommend this unconditionally. These options should be utilized when the medical practitioner is satisfied that their patient is fit to drive (Option 1), or when there is obvious impairment such as moderate to severe dementia or levels of visual acuity that do not meet NZTA minimum requirements (Option 4).

Option 2 has two sub-options. The first is for the medical practitioner to supply a medical-fitness-to-drive certificate with set conditions added to the licence, such as wearing corrective lenses while driving, distance restrictions, or daytime only restrictions. Langford and Koppel (2011) surveyed licence restrictions for drivers aged 65 and over in the Australian state of Victoria. They found that restrictions specific to distance restrictions made up only 0.5% of all those with restrictions, and restrictions for daytime only driving making up 0.3% of restrictions. 96.3% of restrictions were for wearing corrective lenses while driving. They found a trend for those with distance and daytime restrictions, as well as restrictions to drive in specific areas only to have lower crash rates, but these were not statistically significant due to the low base rate of restrictions in the sample. There is currently little evidence that driving restrictions make drivers safer, and restrictions other than for wearing corrective lenses are not regularly used.

The second sub-option according to Figure 2-4 is for the medical practitioner to supply a medical-fitness-to-drive certificate subject to the patient satisfactorily undertaking an On-road Safety Test. If the patient fails that test, this information is forwarded to NZTA who make a decision regarding renewal of the patient's licence. The On-road Safety Test is a 30-min on-road assessment which assesses basic driving skills (e.g., leaving the kerb, turning left at an intersection), hazard detection (e.g., negotiating a crossroad, stopping or giving way at Stop or Give Way signs), and more complex driving situations (e.g., turning right at a crossroad in medium-to-heavy traffic in a 50 km/h zone). Scoring is based on the performance of predetermined manoeuvres, with error scores weighted and combined to give an ultimate pass or fail score.

Option 3 is an intermediate step for when the medical practitioner is unsure whether a diagnosed or suspected medical illness may be affecting driving safety. In this case, the medical practitioner can refer their patient to a medical specialist (e.g., neurologist, geriatrician, psychiatrist, optometrist), or to an occupational therapist (OT) who can perform a medical driving assessment. On receipt of the specialist assessment results, the medical practitioner follows flowchart options 1, 2, or 4. Medical driving assessments are performed by OTs with specialist training in driver assessment. Unlike the On-road Safety Test medical driving assessments do not make use of predetermined lists of errors with weighted scores, but are considered by NZTA to be a more comprehensive assessment of driving ability. The on-road component of an assessment averages 45-min in length and spans a wide range of on-road driving situations (e.g., different speed zones, single- and multi-laned roads, moving from quieter to more busy roads). OTs assess aspects such as awareness of the road and traffic environment, apparent insight into the driving task and any errors that occur, and whether patients are able to compensate for difficulties posed by their medical condition. An OT assessment of driver safety is based on a combination of outcomes of on-road assessment, cognitive and physical evaluations, and any information gathered from a patient including, but not limited to, driving history and frequency, self-imposed limitations, and orientation to time and place.

2.6.4 Summary

While all New Zealand drivers must be medically fit to drive in order to hold a licence, there is no compulsory testing of this (apart from 10-yearly checks of visual acuity and peripheral vision) until age 75. A specific process for licence renewal for older drivers is provided but the official Medical Aspects of Fitness to Drive handbook (New Zealand Transport Agency, 2009) does not provide statistics on the risk that specific medical conditions pose to driving, rather leaving medical practitioners and other health professionals to estimate this for themselves. Perhaps most concerning is the lack of data provided regarding the risk dementia poses to driving, and few guidelines to follow when deciding when a patient diagnosed with dementia may be in need of additional assessment or to have their licence revoked. This information is present in the literature and has been readily used in providing more detailed advice on the management of patients with dementia in other publications (Canadian Medical Association, 2006; Australian and New Zealand Society for Geriatric Medicine, 2010; Iverson et al., 2010).

CHAPTER 3 -

Off-Road Assessment and Prediction of On-Road Driving

3.1 Objectives

The objectives of this chapter are to review the literature for off-road measures that have been used to predict on-road driving. These include cognitive measures, sensory-motor measures, personality measures, self-reported driving behaviour, driving simulators, and computerized assessment measures. The on-road driving assessment itself is then investigated to determine its predictive validity for future real-world negative driving outcome such as crashes. Following this the hypotheses for the research studies are detailed.

3.2 Cognitive Measures

As detailed in Section 2.4.1, normal ageing leads to a slowing of a number of cognitive processes, most notably the speed with which information is processed. As detailed in Sections 2.3.1 and 2.4.2, dementia is characterized by a progressive loss of general cognitive processes and is associated with an increase in negative driving outcomes such as crashes. Cognitive testing shows that even older drivers who do not meet diagnostic criteria for a cognitive disorder and who perform poorly on on-road driving assessments have, on average, poorer scores on cognitive testing (e.g., Wood et al., 2008). However, these measures are not accurate in discriminating which individual older drivers are likely to have problems on the road. Drivers who are not cognitively-impaired are unlikely to be referred for a driving assessment. Naturally, far more research has investigated driving ability in people with cognitive impairment.

Probably because of its brevity, the Mini Mental State Examination (MMSE) (Folstein et al., 1975) was the most commonly used measure in a meta-analysis of 27 studies that used cognitive testing to predict driving ability in drivers with dementia (12 out of 27 studies) (Reger et al., 2004). Korner-Bitensky et al. (2006) reported that 27% of their occupational therapist sample used the MMSE as part of their testing procedure. Unfortunately, the MMSE is not a robust predictor of driving ability, with results too variable between studies to decide on useful cut scores for defining safety (Molnar et al., 2006). This is likely due to the poor sensitivity of the MMSE in detecting the presence of mild cognitive impairment or mild dementia, where many drivers who are unsafe will score in the 'intact' range on this measure (Nasreddine et al., 2005; Smith et al., 2007). Therefore, global measures of cognitive impairment such as the MMSE are likely not specific enough to yield useful prediction of unsafe driving in individuals with dementia.

In the Reger et al. (2004) meta-analysis, the next most commonly used test after the MMSE was the Trail Making Test B (9 studies). Trail Making A was the third most common (6 studies). The majority of cognitive tests investigated in the meta-analysis were used in a single study, making the averaging of effect sizes impossible for many measures. The meta-analysis included studies that had used a variety of dependent outcomes measures: on-road tests, off-road tests (driving simulators or tests of driving knowledge), or caregiver's report of driving ability. When scores on cognitive tests were compared between control and dementia, cognitive tests were significantly different between groups. However, when control groups were excluded, and only those with dementia were compared between driving outcomes, many significant relationships disappeared, but tests of visuospatial function and executive function remained significant outcomes to differentiate safe from unsafe drivers. It is unfortunate that Reger et al.'s meta-analysis included studies with non-on-road measures of driving ability as the dependent measure. The inclusion of studies that used outcome on driving simulators or driving knowledge tests to determine driving ability would not be considered gold-standard measures of driving ability. Nevertheless, a recent narrative review of the dementia and driving literature (Silva et al., 2009) mirrors Reger et al.'s (2004) recommendation that visual attention and executive function tests appear to have the best relationship with driving in people with dementia. Unfortunately the Silva et al. (2009) study also included studies where the dependent outcome measure was simulated driving and informant self-reports of the participant's driving ability rather than on-road assessment outcome. Studies using cognitive tests should instead focus on finding tests that discriminate between drivers with dementia (and do not include control participants) who pass or fail an on-road assessment, or who meet some other criterion that relates to real-world driving ability such as the incidence of crashes or traffic offences.

With the increasing emphasis on the detection of unsafe driving due to the increase in number of older adults over the coming decades there is also a need to investigate the usefulness of cognitive tests in older adults without diagnosed cognitive impairment to determine whether a subset of tests could be used in a primary practice setting to identify older drivers who are at increased risk of negative driving outcomes. As cognitive disorders are prevalent in the older adult population, as discussed in Section 2.3, the detection of at-risk older drivers will increasingly become the task of GPs, and early detection of potential cognitive problems would be useful for the planning of driving changes and possible future cessation of driving. For countries such as New Zealand where GPs are obliged to make a decision about driving safety for their patients aged 75 and over, the availability of an objective and validated screen for the detection of at-risk older drivers would be an invaluable resource.

3.2.1 DriveSafe and DriveAware

DriveSafe and DriveAware are commercially available assessment tools for driving specialist occupational therapists which, when used together, aid in the prediction of on-road driving assessment outcome in drivers referred to driving rehabilitation centres for a variety of medical conditions (http://www.pearsonpsychcorp.com.au/productdetails/374).

The first assessment component, DriveSafe, consists of a series of 13 images of a roundabout with pedestrians and vehicles projected onto a screen (Kay et al., 2009c). The positions and numbers of pedestrians and vehicles vary amongst the 13 images. Examinees view each image for 3 seconds and then report details about the position and direction of travel of the pedestrians and vehicles for each image. The score is the number of items correctly described (maximum of 140 points). Although Kay et al. (2009c) make no claims as to which cognitive processes DriveSafe is measuring, it would appear to be attention, visual scanning, processing speed, and memory.

The second assessment component, DriveAware, was developed as a measure of driving awareness, again for driving rehabilitation centre referrals (Kay et al., 2009a; Kay et al., 2009b). DriveAware consists of eight questions (e.g., why have you been referred for a driving assessment? How would you rate your driving performance now compared with 10 years ago?). Examinees rate each question on a 3-point scale (1 = very aware to 3 = very unaware) and the assessor also rates each answer based on their judgment of the examinees awareness (1 = client has poor performance or knowledge to 2 = client has good performance for knowledge). A total difference score between the opinions of the examinee and examiner is used to determine whether the examinee has intact, partial, or absent awareness. DriveAware is not a cognitive measure per se, but is performed in combination with DriveSafe to classify on-road assessment outcome.

One study has examined the ability of a combination of DriveSafe and DriveAware to predict on-road driving outcome (Kay et al., 2009c). The participants were 115 referrals to two driving rehabilitation centres with an age range of 16 to 95 years (mean age 62.2, 79% male). Around one third had a neurological diagnosis such as a stroke or Parkinson's disease, 26% had dementia including mild cognitive impairment, 24% had orthopaedic or spinal injuries, 6% had acquired brain injury and 10% had miscellaneous other disorders. Both DriveAware and DriveSafe were performed and then each examinee completed an on-road driving assessment where each was rated as pass, conditional pass, intervention, and fail. For the study the pass and conditional pass groups were combined to represent the pass group. Cut points were used for both DriveAware and DriveSafe for trichotomizing on-road outcome. 39% of the sample passed, 31% failed, and 30% received an 'intervention' score. The authors investigated an upper and lower cut point for determining pass and fail scores with the middle group recommended for an on-road driving assessment. 50% of participants were predicted as outright pass and fail, and this result was 90% accurate, meaning that for the entire sample the off-road assessment model correctly classified 45% of on-road pass and fail outcomes. Kay et al. (2009c) state that the purpose of their assessment was to reduce the number of drivers who completed on-road driving assessments, so their results are not comparable to studies that predict a binary outcome (e.g., pass/fail). The authors provide a flow diagram for with cut points for both DriveSafe and DriveAware in order to predict the trichotomous outcome. The authors also mention that the on-road assessors were not blinded to off-road assessment results.

3.3 Sensory-Motor Measures

As discussed in Section 2.4.1, normal ageing brings a number of changes in physical processes such as decreased muscle strength and endurance, slower reaction and movement times of limbs, and reduced physical mobility, particularly in the trunk, neck and head. The research literature on driving has tended to focus on cognitive changes associated with decreased driving ability, with less attention paid to the effects of sensory-motor changes. A narrative review by Stelmach and Nahom (1992) investigated sensory-motor decline in older adults and their link with possible changes in driving ability and suggested that motor control is of prime importance in situations where emergency manoeuvres are required, with older drivers likely to find it more difficult to plan and execute these tasks than younger drivers. The authors split their review into two broad sections of reaction time and movement time. Reaction time measures include the cognitive tasks required to plan and initiate a movement (i.e., response preparation, response selection, response programming, etc.), with movement time referring to the actual process of motor activity (i.e., acceleration and deceleration, coordination, and joint flexibility). The review investigated cross-sectional laboratory studies that quantified the differences between young adults and older adults and found increasing difficulties with these tasks as age increased.

In terms of sensory-motor measures, Wood et al. (2008) found that postural sway, which had previously been found to be related to the risk of falling in older people, was a significant predictor of on-road driving score in a sample of cognitively-unimpaired drivers aged 70 and over. Innes et al. (2007) found that a computerized measure of tracking a moving line target with a steering wheel and reaction time for turning a steering wheel following presentation of a visual stimulus were predictors of on-road driving ability in a group of drivers with brain disorders (mostly stroke). Stav et al. (2008) found that the motor measure with the largest correlation with on-road assessment outcome for a mixed group of impaired and unimpaired drivers aged 65 and over was the time to walk ten feet and to turn around and walk back again. This measure was also retained in their final model for classifying on-road outcome alongside cognitive tests.

Sensory-motor measures appear to have a relationship to driving ability in both cognitively-impaired and unimpaired older drivers. Hence these measures could add predictive power over and above that afforded by cognitive tests and could be particularly useful in cognitively-unimpaired older drivers who are none-the-less subject to the process of normal ageing of the body.

3.4 Personality Measures

Relatively little research has questioned whether personality variables influence driving safety, and this research has primarily focused on young adult samples. A study with a mean sample age of 39 years found emotional stability, accepted level of risk, and social responsibility to be significant classifiers of on-road driving assessment outcome (Sommer et al., 2008). In college students, low scores on the personality construct of conscientiousness and high scores on sensation-seeking have been associated with higher rates of both self-reported and simulated risky driving behaviour (Schwebel et al., 2006). Schwebel et al. (2007) later examined drivers aged 75 and over and found that higher scores on sensation-seeking were correlated with self-reports of higher numbers of driving violations and tickets (correlations ranging in size from .24 to .30). Unfortunately, no sensation-seeking scales specifically constructed for, or validated against, older adult samples could be found, which brings into question whether scales constructed for younger adults are valid to use in older adult samples.

The propensity to become angry whilst driving is another personality trait that may affect driving safety. The Driving Anger Scale (DAS) is a self-report scale developed by Deffenbacher, Oetting, and Lynch (1994). Higher scores on the 14-item short form of the DAS have been associated with increased self-reported risky driving behaviour in young adult samples (Deffenbacher et al., 2002; Deffenbacher et al., 2003; Dahlen & White, 2006). However, the relationship between DAS scores and negative driving outcomes in older drivers has not been investigated.

3.5 Self-Reports of Driving Behaviour

Reason et al. (1990) developed the Driver Behaviour Questionnaire (DBQ) to investigate whether reporting specific instances of mistakes and road rule violations on the road was able to identify drivers at higher risk for adverse driving events. This questionnaire asks about the frequency of occurrence of 24 driving behaviours over the previous 12 months. The behaviours fall into three subscales of lapses of attention, errors, and deliberate violation of road rules. Several studies have used this scale as a predictor of on-road driving. Endorsement of a greater number of items of the violation subscale has been associated with increased crash involvement in young adult samples (Parker et al., 1995; Sullman et al., 2002). However, older drivers report fewer violations than younger drivers (Davey et al., 2007). Parker et al. (2000) found a positive association between increased lapses and errors and a higher crash rate in drivers aged 50 and over. These authors suggest the scale may be useful in detecting declining cognition which leads to lapses in concentration indicative of a decline in driving safety, although further study is required to replicate this finding.

3.6 Driving Simulators

Driving simulators seek to simulate a realistic and face-valid driving environment. The driving simulator environment has advantages over real-world driving in that specific driving situations can be presented that would be impossible in an on-road assessment, such as programming pedestrians or other vehicles to cross the path of the driver, or pre-determining the traffic density encountered during a drive. A major obstacle with the use of simulators is determining the extent of the relationship between errors performed on the simulation and errors performed in real-world driving situations. This is often not investigated yet is essential knowledge if simulators are to be used as a proxy for real-world driving assessments. Simulators can also be very expensive to purchase, with a 2004 New Zealand report (Bowens, 2004) stating the price for the System Technology Inc. (USA) STISIM Drive™ three-monitor driving simulator at US$59,000. Pricing is thus a restriction for the widespread use of advanced driving simulators as routine measures of driving ability. Feelings of dizziness or ‘simulator sickness’ are also a concern, with 9% of one older adult sample experiencing feelings of dizziness following completion of a simulated drive (Lee, Lee, & Cameron, 2003). Another study reported simulator sickness in 17% of participants, with older drivers more often affected than young (Brooks et al., 2010). This section does not purport to be detailed review of driving simulators, and presents research on three systems, as these focused on predicting driving ability in drivers with a neurological disorder: the Systems Technology Inc, Simulator (STISIM Drive™), DriVR™, and Virtual Reality Driving Simulator (VR-DR).

3.6.1 STISIM Drive™

The STISIM Drive (http://www.stisimdrive.com/) system consists of driving controls including a dashboard with steering wheel and indicator stalk, accelerator and brake pedals and one or more computer screens that display a realistic driving scenario which includes buildings, other vehicles, and pedestrians in a range of programmable driving scenarios (Systems Technology, 2010). The software is able to be run in a number of different hardware configurations including a scale-model vehicle cab, a simulated vehicle cab, as part of a modular desktop arrangement, or as a portable device using a laptop computer (Figure 3-1).

Figure 3-1. The STISIM Drive™ Training Simulator model M400 which includes three monitors which present a 135° driver field-of-view. Taken from http://www.stisimdrive.com/images/stisim/ content/ downloads/m400.pdf.

The STISIM Drive has been used in a number of driving research projects, mainly conducted by Lee and colleagues at the School of Occupational Therapy at Curtin University of Technology, Perth, Australia. Using STISIM Drive with a single computer monitor screen, Lee, Lee, and Cameron (2003) found a moderate positive correlation ($r = 0.51$) between age and reaction times in response to an on-road stimulus in 129 cognitively healthy drivers aged 60-89 (mean age 72.9 years, 78% male). Using the same sample, Lee, Cameron, and Lee (2003) found that performance on STISIM explained 65.7% of the variance in on-road assessment performance. Further investigation of this same sample found that several measures recorded using the STISIM were related to experiencing a self-reported crash over the previous 12 months (a one point increase in scores on the working memory, decision and judgment and speed compliance components were associated with decreased risk of a crash of 45%, 61% and 17% respectively) (Lee, Lee, Cameron et al., 2003). Two years following this assessment those who had received poorer scores in simulator indicator use and better scores on the working memory component had higher rates of traffic violations over this period (23% and 16% increases respectively) (Lee & Lee, 2005). The authors claim that a better score on the working memory component is a logical outcome "because drivers with good working memory and confidence tend to drive above the speed limit" (Lee & Lee, 2005, p. 100). An independent pilot study of only 9 participants aged 67-78 found a significant correlation ($r = -0.83$) between lower error scores on the STISIM and failing an on-road driving assessment (Freund et al., 2002). Another study reported that the performance of many errors performed in an on-road driving assessment did not differ significantly from those performed in a STISIM Drive course (Shechtman et al., 2009). In order to determine the predictive validity of the STISIM Drive, large-scale studies comparing performance on the stimulator to real-world outcomes such as crashes, traffic offences, or on-road driving ability are required.

3.6.2 DriVR™

DriVR is a PC-based system formerly marketed by Imago Systems of Vancouver, Canada. DriVR incorporates a steering wheel, foot pedals, and virtual-reality glasses (Liu et al., 1999). A virtual environment is presented through the glasses with a 30-degree horizontal field of view that allows examinees to rotate their heads to scan left and right in the virtual environment. DriVR depicts driving scenarios with a number of programmed driving situations, such as driving down a street and pulling into a driveway. A number of quantitative measures are taken during an assessment, such as speed of travel, position within the lane, and avoidance of crashes. Qualitative measures can also be taken by the examiner, such as the number of head movements made to the left and right during the testing session.

One study of 162 people aged from below 16 to above the age of 76 found that older drivers drove more slowly than younger participants and stopped at stop signs less frequently (Liu et al., 1999). Liu et al. (1999) also compared the performance of 17 people with brain injuries to age and sex-matched controls. They found the brain-injured group had significantly poorer scores on several DriVR measures compared to the controls. Only one study compared performance on the DriVR to real-world driving as measured by an on-road driving assessment (Wald et al., 2000). In this study, 28 participants with brain injury with a mean age of 40 years completed the DriVR course as well as two on-road assessments. Wald et al. (2000) found very small- to moderate-sized correlations (values ranged from 0.01 to 0.56) between several DriVR scores and various on-road driving scores in drivers aged 20-76 (mean age of 40 years). The authors did not, however, provide *p* values for these associations, so it is unknown if any of the correlations were significant. There has been no research published on the DriVR since 2000, and the DriVR website no longer exists which suggests that the simulator may no longer be available.

3.6.3 Virtual Reality Driving Simulator (VR-DR)

The VR-DR is a research driving simulator based in the Psychology Department of Drexel University, Philadelphia, and has been used in the study of the effects of brain injury on driving tasks (Schultheis & Mourant, 2001). The simulator consists of a head-mounted display unit that presents a computer-generated virtual-reality environment through which a participant drives using a steering wheel and foot pedals (see Figure 3-2). The head-mounted unit displays a 60° field of view at any one time which updates with head tracking to present a full 360° field of view.

Figure 3-2. The VR-DR simulator in use. Adapted from Schultheis et al. (2007).

The simulated drive is around 30 minutes in length and presents a variety of different driving scenarios for the examinee to navigate as well as challenges such as avoiding a pedestrian who suddenly crosses the road. A study examined user ratings of the VR-DR in 33 participants with acquired brain injury (61% had a moderate to severe TBI with the remainder having had a stroke) and 21 healthy controls, all aged less than 68 years with a mean age of 41.7 years (Schultheis et al., 2007). The brain-injured participants provided a lower overall rating for the ease of use of the VR-DR. Another study looked at reactions related to stopping at stop signs in a sample of 15 adults with brain injuries (67% with moderate to severe TBI and the remainder with stroke) and 9 healthy controls (Schultheis et al., 2006). Both groups showed improvements in appropriate stopping procedures (such as the distance from the sign and coming to a complete stop), with a trend for the brain injured group to perform slightly less well than the controls. No published studies have compared the VR-DR to real-world driving through driving records or on-road assessment ratings, and none have investigated its use with older drivers or drivers with cognitive impairment due to MCI or dementia.

3.7 Computerized Assessment Systems

Computerized driving assessments seek to measure the underlying sensory-motor and/or cognitive processes that mediate a person's ability to drive but, unlike driving simulators, they do not purport to create a realistic visual representation of a driving environment. This review focuses on four of the more commonly reported off-road computerized systems: the Elemental Driving Simulator, the Useful Field of View®, DriveABLE™, and Sensory-motor and Cognitive Tests™.

3.7.1 Elemental Driving Simulator (EDS)

The EDS was developed by Gianutsos (1994) and took over from its forerunner the Driving Advisement System (Gianutsos et al., 1992). The EDS focuses on the cognitive factors that affect driving safety in people who have experienced brain injury. The hardware of the EDS consists of a computer with monitor, steering wheel, turning indicator stalk mounted on the steering wheel, and foot pedals (see Figure 3-3). Software runs in the DOS operating system.

The EDS testing protocol begins with the examinee asked to self-rate how well they think they will perform a number of driving-related tasks. This is followed by three subtests that measure pursuit tracking with the second and third tests adding a measure of reaction time and then choice reaction times respectively (the indicator stick is move up or down to select the side of the screen where a face shaped stimulus is flashed). The most recent version of the EDS was released in 2004.

The developers of the system compared performance on the EDS in three different participant groups: 50 'normative' drivers (mean age 41 years), 1,145 older drivers from the community (mean age 69 years), and 85 people referred for driving rehabilitation due to brain injury, stroke, or developmental neurological conditions (mean age 37 years) (Gianutsos, 1994). Unfortunately, only the sample referred for driving assessment completed an on-road assessment, where each was rated as pass or fail.

Figure 3-3. The EDS Portable Inclusive System showing computer, steering wheel with indicator stalk to the left, and foot pedals.

Results on the EDS system are presented by the authors only in graphical format with the ‘normative’ drivers showing the best scores for the four groups. Gianutsos notes that all members of this group completed the EDS assessment but that not all members in the other groups did, although the numbers and reasons for non-completion are not reported. Referrals who passed the on-road test generally performed better across measures than the group who failed, although the sample of older drivers did nearly as poorly as the referral grouped who failed the on-road assessment on all but one measure. Gianutsos claims the EDS is a useful system for detecting those with unsafe driving, but this cannot be substantiated since only the neurological group completed an on-road driving assessment. Also, the scores on the EDS were very similar between the older driver group and the neurological disorder group who went on to fail an on-road assessment, suggesting that the EDS may over-report that older drivers are unsafe.

An independent Australian and New Zealand joint study of driving assessment for drivers aged 80 and over reported that 16% of their sample found the EDS too difficult to complete (Monash University Accident Research Centre, 2004). There were also no significant associations found between the outcome measures of the EDS and on-road driving performance. These outcomes, as well as technical difficulties with the system, led to the

EDS being dropped from further analysis for the study. On the basis of this evidence, the EDS would not be considered a useful screen for determining whether an individual is likely to pass or fail an on-road driving assessment.

3.7.2 Useful Field of View® (UFOV)

The UFOV was designed to detect the capacity of a person's preattentive processing (Owsley et al., 1991). Preattention processing is described as working at a parallel (versus serial) level and is designed to capture and direct a person's attention to salient visual stimuli. The UFOV is different from visual field as measured by clinical perimetry but is dependent on the visual field in order for stimuli to be perceived. The authors devised three subtests to assess different mechanisms that can each restrict a person's useful field of view: slowing of information processing, impaired ability to divide attention, and impaired ability to ignore visual distractors. The three subtests all require that the examinee fixate their vision on a central stimulus presented on a computer screen and make a binary decision about that stimulus (i.e., whether two presented symbols match or not), with the first test being a simple reaction time test that is used as a baseline measure for the following two tests. The divided attention subtest requires that the examinee simultaneously report the location of another stimulus on the periphery of the screen while completing the central task. The distractors test further clutters the visual array with distracting stimuli while the other tasks are performed.

Lower scores on the UFOV have been associated with an increased likelihood of having a crash in older drivers, both retrospective to the assessment (Owsley et al., 1991; Ball et al., 1993; Sims et al., 1998; Ball et al., 2006) and in the three years following it (Owsley et al., 1998). However, two other studies did not find the UFOV useful in classifying those who had self-reported crashes compared to those who had not (De Raedt & Ponjaert-Kristoffersen, 2000; Hoffman & McDowd, 2010). Two studies found the UFOV a useful predictor of prospective driving cessation in older drivers over three years (Ackerman et al., 2008) and ten years (Edwards, Bart et al., 2009). The UFOV has also been investigated in terms of on-road driving assessment outcome, with three studies finding it was a significant predictor (in addition to other measures within the same model) to classify those older drivers who went on to perform more errors in an on-road assessment (De Raedt & Ponjaert-Kristoffersen, 2000; Whelihan et al., 2005; Uc et al., 2009), and one study finding it useful in classifying on-road fails (Stav et al., 2008). One study, however, did not find the UFOV useful in classifying on-road driving errors in a group of drivers with Alzheimer's dementia (Dawson et al., 2009), and another study did not find it useful for predicting driving score in a sample of healthy drivers aged 70 and over (Wood et al., 2008).

Duchek et al. (1998) found that people with a Clinical Dementia Rating of 1 (mild dementia) had difficulty completing the UFOV, and suggested that a simplified version may be needed for this population. Another problem is that Ball et al. (1988) found that participants could improve their performance on the UFOV with practice, and that the effects lasted as long as six months. No studies have investigated whether training on the UFOV corresponds with safer driving practice.

Substantial research on the UFOV suggests it could be a useful measure for detecting people with possible driving impairments, with several studies finding that the divided attention subtest (subtest 2) is the most useful UFOV measure (Owsley et al., 1998; Ball et al., 2006; Edwards, Bart et al., 2009). However, its difficulty with persons with dementia is an impediment to its use for higher risk groups of older drivers who likely contain a substantial percentage of people with cognitive disorders.

3.7.3 DriveABLE™

DriveABLE Assessment Centres Inc. is a Canadian company founded by Allen Dobbs following research to develop an off- and on-road driving assessment protocol for medically-impaired drivers (Dobbs, 1997). The DriveABLE assessment consists of both off- and on-road testing components. The off-road component is called the DriveABLE Cognitive Assessment Tool, which consists of a series of tests of memory, attention, reaction times, and judgment which are performed using a touch sensitive computer screen or by pushing a button (Dobbs, 2005; *DriveABLE*™, 2011). Raw test results are sent electronically to the DriveABLE centre in Alberta, Canada, with performance scored and a predicted probability of failing the on-road assessment sent back, along with a predicted trichotomous rating of pass, fail, or indeterminate for on-road performance. Only those with an indeterminate rating are recommended to complete the standardized on-road test. The results of the on-road test are also scored according to a DriveABLE algorithm, with a possible outcome of recommended pass, borderline pass, and recommended driving cessation returned to the assessor. DriveABLE is currently used in 76 locations, mostly in the US and Canada, with one assessment centre each in Australia and New Zealand (*DriveABLE*™, 2011).

The algorithms for determining outcome scores were based on the recruitment and testing of a group of drivers with dementia (n=176) and two control groups, one of drivers aged 55 and over (n=70), and one of younger drivers (n=33) (Dobbs, 1997, 2005). Errors performed on-road were compared between groups, with the dementia group defined a-priori as 'unsafe' due to the increased driving problems found in people with dementia, and the other two groups defined as safe. Researchers then developed a list of errors present primarily in the dementia group and defined these as the errors both off- and on-road that would be used to determine driving safety. While the results of the initial and validation testing have not been published by Dobbs, a summary can be found in a U.S. Department of Transportation report (National Highway Traffic Safety Administration, 1999). This report shows that 431 drivers were recruited for the validation study. 67% of the sample were assigned a rating of pass or fail for the off-road assessment, with the remaining one third of the sample recommended for on-road assessment (presumably the on-road assessors were blinded to off-road testing results, but this information is not provided). The report says that the screen was 94% accurate in classifying pass and fail outcomes, but it should be noted that since only 67% of the sample were given a pass or fail rating the actual classification rate is 94% of 67%, or 63% overall accuracy of classification.

Korner-Bitensky and Sofer (2009) found that 67% of a sample of 52 drivers referred for medical driving assessments were correctly predicted as 'recommend cessation', 'indeterminate' , and 'no evidence of reduced competence' by the DriveABLE off-road assessment relative to blinded DriveABLE on-road assessment scores of pass, borderline pass, or fail. The authors then dichotomized their off-road results into those who had a predicted fail, and those who had an indeterminate score or pass recommendation, and dichotomized their on-road results into fail, and borderline pass and pass outcomes. They found sensitivity for detecting fails of 76% (32 of 42 on-road fails recommended as fails in the off-road assessment) and a specificity of 90%. The positive predictive value was 97% (of the 33 participants for whom cessation was recommended, 32 failed the on-road test) and a negative predictive value of 47% (of the 19 participants who were classified as 'indeterminate' or 'no evidence of reduced competency' nine passed the on-road assessment). The authors note these positive and negative predictive values reflect the rate of fails and passes in the sample, and the high rates of fail scores in their sample (81%) would have affected these results. They recommended that others replicate their study to determine the stability of these values.

One other major study of the DriveABLE off-road prediction was performed for the Australian and New Zealand Austroads study (Monash University Accident Research Centre, 2004). This study did not use the DriveABLE on-road assessment but used the predicted probability of off-road outcome (a score between 0.0 and 1.0 for each participant) relative to pass or fail outcomes on their own on-road driving assessment. When classifying their pass and fail on-road outcomes and using the cut-point with the highest average of sensitivity and specificity, this study found a sensitivity for classifying fails of 65.3%, a specificity of 57.9%, a positive predictive value of 35.6% and a negative predictive value of 82.4%. This study had a much lower fail rate than the Korner-Bitensky and Sofer study (2009) (36%) and comprised 300 cognitively-healthy older drivers aged 80 and over. These are likely two reasons for the lower classification in this group.

DriveABLE is a difficult system with which to perform an independent study as it is reliant on scoring by the DriveABLE company and subtest scores cannot be examined for their individual relationships to on-road outcome. There is little detailed published research by the developer, and measures of sensitivity and specificity quoted are inflated due to exclusion of the results of those rated as indeterminate on the off-road assessment. Perhaps the primary limiting factor for DriveABLE is the automatic rating of drivers with dementia in the training data set as unsafe and the implicit acceptance of all errors performed by non-demented driver as safe. This is in contrast to recruiting a mixed sample of drivers with and without cognitive impairment and determination of the errors which are related to unsafe on-road performance on the basis of on-road outcome, whether that includes cognitively-unimpaired drivers or not. Dobbs (2005) explicitly states that only older drivers with medical conditions should be considered for driver assessment, and has been careful to define the differences in driving errors between cognitively-impaired and unimpaired older drivers (Dobbs et al., 1998; Dobbs et al., 2004). Essentially, the algorithm is weighted to penalize drivers with cognitive impairment and to pass those without. Thus, the prediction is not as useful for detecting *which* drivers with cognitive impairment are safe and which are not, unless the aim is to only pass those safe drivers with cognitive impairment whose driving behaviour is indistinguishable from those drivers without cognitive impairment. This would effectively mean the removal of licences for all people exhibiting the first signs of cognitive impairment due to dementia, stroke, head injury, and other causes.

3.7.4 Sensory-motor and Cognitive Tests™ (SMCTests™)

SMCTests™ is a battery of sensory-motor and cognitive tests developed by researchers from Christchurch and Burwood Hospitals, the Van der Veer Institute for Parkinson's & Brain Research, University of Otago, and University of Canterbury as a clinical research tool with application in the assessment of patients with neurological disorders and in ageing (Christchurch Neurotechnology Research Programme, 2006; Innes, Jones, Anderson et al., 2009). *SMCTests* has been tailored into the semi-portable Canterbury Driving Assessment Tool (*CanDAT*™) which comprises a laptop computer, a separate colour screen to display test stimuli, a steering wheel with two indicator levers (one on the left and one on the right), and a set of pedals with accelerator and brake, and an optional clutch pedal. Tests include measures of visuospatial function (Jones & Donaldson, 1995), visuoperception (Jones & Donaldson, 1995), reaction times of arms and legs, ballistic movement of arms, tracking (Jones & Donaldson, 1986; Jones, 2006), decision-making, visual search, complex attention, impulse control, and planning (Innes et al., 2007; Innes, Jones, Anderson et al., 2009). The full battery takes around an hour to complete depending on the level of impairment. The *SMCTests* battery is currently used in several occupational therapy practices in New Zealand as part of medical driving assessments, mostly in the form of the *CanDAT*, but with one centre using the system in a modified car body (Figure 3-4).

Figure 3-4. The modified car body used to run *SMCTests* pictured left. Test stimuli are projected onto the white wall in front of the machine. The semi-portable *CanDAT* apparatus is pictured right.

A clinical study investigating driving ability in 501 people with brain disorders used *SMCTests* to correctly predict 76% of referrals as on-road Pass or Fail relative to a blinded on-road driving assessment carried out at the Driving and Vehicle Assessment Service (DAVAS) at Burwood Hospital (Innes et al., in press). Compared to the other computerized tests described above, *SMCTests* provides on face value a more detailed account of sensory-motor function in examinees. While DriveABLE can measure reaction times, it cannot assess movement time and velocity of arm movements. The EDS systems utilizes tracking using a steering wheel, but also does not measure velocity and movements times of the arms. The UFOV also measures reaction times, but again does not measure velocity of arm or leg movements.

3.8 On-Road Assessment and Real-World Driving Outcomes

The rationale behind on-road driving assessments for experienced drivers is that behaviours observed during the assessment are indicative of driving behaviours that are performed in real-world driving. That is, a person who appears to have a lack of awareness for other road users is assumed to have this problem when driving on their own and that it puts them at increased risk for negative driving outcomes. While this may seen face evident, there is little evidence to substantiate this claim as drivers who fail on-road driving assessments are usually prevented from driving and, hence, unable to be followed to determine if they have more negative driving outcomes than drivers who pass on-road assessments.

Only one study was found that prospectively followed the crash outcomes of a group of ostensibly healthy older drivers following *both* pass and fail outcomes on an on-road test. An Australian study followed 266 older drivers from a population-based sample for 12 months following an on-road driving assessment, with participants providing monthly self-reports of crashes over this period (Anstey et al., 2009). The authors found no increase in the incidence of self-reported crashes for drivers who had received a score in the Fail range of an on-road driving assessment. However, only 6% of drivers scored in the Fail range on the on-road assessment which would have limited the power for finding a significant increase in crashes for the Fail group.

A New Zealand study prospectively followed the incidence of police-reported injury crashes in a population sample of over 39,300 drivers aged 80 and over following pass scores obtained on an on-road driving test (Keall & Frith, 2004a). The test was administered under driver licensing laws (now defunct) that required adults aged 80 and over to pass biennial on-road driving tests. There was no limit to the number of times drivers could sit the test in order to secure a passing grade which provided a rare opportunity to determine whether drivers who failed one or more on-road assessments (but who eventually passed) were more likely to experience adverse driving events in the following two years. Seventeen percent of the sample failed the on-road test at least once before receiving a pass grade and the risk of involvement in a crash in the following two years rose 33% for each time the test was re-sat. Unfortunately, the study tells us nothing about the 5% of drivers who never passed the on-road assessment and therefore did not have their licences renewed. The study also had no information about the presence of cognitive impairment and other medical problems in the sample.

Data provided in the Keall and Frith study was sufficient for a phi correlation coefficient to be calculated to determine the effect size difference in rates of crashes of the on-road assessment of the group with at least one fail compared to the pass group. To calculate the phi coefficient, the study data were dichotomized into those drivers who passed their on-road test on the first administration (32,358 people) versus those who sat two or more on-road tests before receiving a passing grade (6,943 people) with rates of serious crashes compared between the two groups, see Table 3-1.

Table 3-1. Number of people in Keall and Frith's (2004a) study divided into those who passed and failed their first on-road test and those who went on to have a Police-recorded crash over the following 24 months

	24-month police-recorded crash involvement			
First on-road test	No	Yes	Totals	% who had a crash
Fail	6863	80	6943	1.2%
Pass	32135	222	32357	0.7%
Totals	38998	302	39300	0.8%

Drivers who failed one or more tests were significantly more likely to have a crash in the following 24-months than drivers who passed on the first test administration (Fisher's Exact Test two-tailed, $p < .001$). The phi coefficient, however, was very small at $r_\varphi = 0.02$. Due to their extremely large sample size, Keall and Frith were able to find a significant, but very small, association between failing on-road tests and later crashes.

A major problem with on-road driving assessments is that most driving assessors do not use standardized scoring and do not use pre-determined cut scores for assigning pass or fail ratings. Korner-Bitensky et al. (2006) surveyed the driving assessment methods of 144 American and Canadian driving assessors. Ninety-four percent of respondents routinely used on-road assessments as part of their evaluation yet only 24% used a standardized scoring system and only 10% used a pre-defined cutoff score to define driving competency. Only two respondents reported using a standardized road test. Some standardized assessments have been tested for inter-rater and test-retest reliability, with the former usually found to be moderate to high, and the latter in the moderate range (Hagge, 1994; Fitten et al., 1995; Romanowicz & Hagge, 1995; Hunt et al., 1997; Janke & Eberhard, 1998). Investigations into the validity of standardized road tests have found some associations to real-world crashes or infringements (Fitten et al., 1995; Romanowicz & Hagge, 1995; Keall & Frith, 2004a), although due to the low base rates of crashes in particular, power is low for detecting statistically significant associations. Other methods to test on-road assessment validity have been based on finding differences in group performance in expected directions, such as differences in error scores or Pass and Fail results between novice and experienced drivers (Hagge, 1994; Romanowicz & Hagge, 1995).

A major reason for not using standardized scoring is the flexibility that a non-standardized format allows an assessor in drawing on their often considerable clinical experience in deciding whether a person is a safe driver. For example, a standardized route cannot be used when an assessment is performed from a person's home and in their local area because they only use their car once a week to do the grocery shopping and to attend doctor's appointments. A standardized scoring system also misses the full range of factors influencing the outcome of an assessment. For example, missed appointments may indicate a memory problem, and the impact of confusion may be a strong indicator of unsafe driving that cannot be simply expressed in a checkbox rating. This flexibility may be particularly important when assessing people with cognitive impairment or dementia due to the idiosyncratic way impairments affect the abilities of individuals. We believe, however, that it is important that at least some standardization takes place in order that judgments made between drivers and between assessors are as equitable as possible.

3.9 Summary

Researchers have explored a range of methods to measure the cognitive and sensory-motor processes that underlie driving, and to find measures that discriminate between those judged as safe or unsafe drivers. Cognitive measures alone have not produced models strong enough to make judgments of driving safety based on their outcomes alone, perhaps because the over-learned procedural task of driving does not correlate well with specific measures of cognitive domains. There are also some promising sensory-motor measures, such as the rapid pace walk, that could be assessed in future studies. There is little current evidence that personality measures could be useful to predict driving in older adults, but their ease of inclusion in studies would recommend them as potentially useful measures for future research.

Driving simulators have advantages over real-world on-road testing of being able to present driving situations that cannot be performed in real-world driving scenarios. Further research is needed, however, to demonstrate their ability to differentiate between drivers who would be considered safe and unsafe either through on-road driving assessment or through records of crashes and traffic offences. Computerized assessments require the same kind of studies required of simulators in order to show that their outcomes are related to real-world driving, and both simulators and computerized assessments need to be more thoroughly tested with cognitively-impaired and unimpaired older drivers to determine if a reasonable percentage of examinees are able to understand and complete the assessments. As with all off-road assessments, driving simulators and computerized tests are safe as they do not involve exposure to risk through actual on-road driving.

In order to find the best predictors of driving ability, it is necessary to include a broad range of possible predictors from multiple domains. It could be that a combination of measures from different domains provide increased predictive power over focusing on a single domain such as cognition or personality. A combination of cognitive measures, sensory-motor measures, and personality factors could provide a more comprehensive understanding of driving than sticking to single domains, as all these measures have shown some degree of relationship to on-road outcomes.

CHAPTER 4 - Study Hypotheses

The primary aim of the three studies described in this thesis was to find predictors of the ability to Pass or Fail an on-road driving assessment in (1) a group of cognitively-unimpaired older drivers and (2) a group of drivers with mild cognitive impairment or Alzheimer's dementia. Given that cognitively-unimpaired drivers in the first study would suffer no consequences for failing the on-road assessment we were in the position to be able to follow this group of drivers to determine whether there was any relationship between measures taken at the baseline assessment or the Pass/Fail outcome of the assessment and whether participants went on to experience negative driving outcomes over the following few years.

The hypotheses for each of the three studies are presented below and follow on from the literature reviews in Chapter 2 and Chapter 3.

4.1 Healthy Older Drivers Study

The primary aim of the Healthy Older Drivers study was to measure the association between off-road measures of cognitive testing, personality factors, and sensory-motor testing and on-road assessment Pass and Fail outcomes. Following this, measures most associated with Pass and Fail outcome were to be offered to a binary logistic regression model in order to find the most parsimonious group of measures to classify on-road outcome. The classification model formed was then to be tested using leave-one-out cross-validation to determine the stability of the model and how likely it would be to generalize to a new sample. In addition to these major goals of the study, specific hypotheses were investigated.

4.1.1 Hypothesis 1

Gap in the Literature: No previous study had investigated the ability of *SMCTests* to aid in the prediction of on-road driving assessment outcome in cognitively-unimpaired older adults. There are also no other computerized tests of driving ability that encompass the breadth of sensory-motor testing offered by *SMCTests*.

Hypothesis: A combination of standard cognitive tests and *SMCTests* measures will provide on-road Pass and Fail accuracy statistics higher than that achieved by previous studies.

Rationale: All older drivers are affected by the physical changes which occur as part of ageing (as described in Section 2.4.1). Most studies of older drivers concentrate on cognitive changes that could predict unsafe driving, with few having sensory-motor measures other than simple reaction time of the upper limb or limbs. *SMCTests* includes a variety of reaction and movement time measures of both upper and lower limbs, as well as measures of maximum velocity achieved. These additional visuomotor, coordination, and visuospatial sensory-motor measures in combination with standard and computerized cognitive measures may provide an off-road driving assessment with a higher accuracy for classifying participants into on-road assessment Pass and Fail groups, particularly for older adults without manifest cognitive impairment but who are still subject to the physical changes associated with ageing.

Significance: Finding a combination of tests that predict driving ability in cognitively-unimpaired older drivers would allow for more certainty in the recommendation of continued driving currently performed by general practitioners. A screen for possible impairments could alert general practitioners to drivers who could benefit from a comprehensive driving assessment. This would lead to fewer unsafe older drivers on the road.

Study design (Chapter 6): 60 drivers aged 70 and over will be recruited to complete a selection of standard cognitive tests and a subset of *SMCTests* measures. All participants will complete an on-road driving assessment administered by a driving specialist occupational therapist and a driving instructor, and will be assigned a Pass or Fail rating and a score on a 0-10 driving scale. A binary logistic regression model will be formed to classify participants into Pass and Fail groups, with the model then tested using leave-one-out cross-validation to simulate the stability and predictive power of the model in an independent test set.

4.1.2 Hypothesis 2

Gap in the Literature: No previous studies have investigated the ability of the Driving Anger Scale to predict on-road driving ability in a group of cognitively-unimpaired older drivers.

Hypothesis: Participants with higher scores on the Driving Anger Scale are more likely to Fail an on-road driving assessment.

Rationale: Higher levels of driving anger have been associated with increased negative driving outcomes in younger adults (see Section 3.4) although the scale has not been used in a sample of older adults.

Significance: If higher scores on the Driving Anger Scale are associated with poorer scores on the on-road driving assessment then the scale could be a potentially useful screening tool for determining which cognitively-healthy older drivers may be at risk for negative driving events.

Study design (Chapter 6): Same as for Healthy Older Drivers Hypothesis 1.

4.2 Healthy Driver Follow-Up Study

The primary aims of the Healthy Driver Follow-up study were to measure the association between on-road Pass and Fail outcomes and scores on initial testing conducted for the Healthy Older Drivers study and subsequent rates of crashes and traffic offences. We were also interested in changes in driving behaviour over the two-year follow-up, and whether aspects of medical conditions were related to crashes and offences. These goals along with some additional hypotheses are listed below.

4.2.1 Hypothesis 3

Gap in the Literature: Only two previous studies (Keall & Frith, 2004a; Anstey et al., 2009) have followed the real-world driving of older drivers following an on-road driving assessment as discussed in Section 3.8. Only one study (Anstey et al., 2009) followed people who had failed the assessment. In both studies the dependent variable was crashes, which are a low base-rate driving outcome.

Hypothesis: Drivers who fail an on-road driving assessment will have a greater incidence of crashes and traffic offences over the following years than those who pass.

Rationale: Drivers who fail an on-road assessment have necessarily demonstrated driving behaviours which are considered by the assessor to be dangerous and likely to mean that the examinee will go on to have negative driving outcomes in real-world driving. We would expect then that drivers who fail the assessment would have more negative driving outcomes than drivers who passed.

Significance: Due to the few studies that have been able to follow drivers who have failed an on-road assessment, a finding that fail scores are associated with an increase in negative driving outcomes would go some way towards validating the use of on-road driving assessments in cognitively-unimpaired older drivers. A positive association between fail scores and subsequent crashes and traffic infringements would demonstrate that even cognitively-unimpaired older drivers can display driving behaviours indicative of future adverse driving events.

Study design (Chapter 7): Drivers in the Healthy Older Driver study will be followed for two years with annual interviews regarding their driving behaviour and self-reported history of crashes and traffic offences over the previous 12 months. Official records of crashes and traffic offences will also be obtained, with self- or officially-reported crashes and offences making up the binary variable of crash or offence or no crash or offence.

4.2.2 Hypothesis 4

Gap in the Literature: Previous research has found various cognitive measures that are related to driving outcomes but few studies have followed participants prospectively over time to see if these measures predict real-world driving outcomes.

Hypothesis: Poorer performance on standard cognitive measures are associated with increased crashes and traffic offences over the following years.

Rationale: Poorer performance on cognitive tests could indicate the presence of as yet undetected cognitive impairment. LaFont et al. (2008) found a relationship between eventual dementia diagnosis and self-reported crashes up to seven years before the diagnosis. This indicates that cognitive changes may impact driving behaviour even before the impairment is great enough to warrant diagnosis.

Significance: Finding a cognitive measure or measures that can predict real-world driving outcomes over a prospective two-year period would be useful for inclusion in screening for driver safety. An indicator of possible real-world driving outcomes coupled with an indication of reduced likelihood of passing an on-road assessment could be a powerful indicator that more comprehensive driving assessment is required.

Study design (Chapter 7): Same as for Healthy Driver Follow-Up Hypothesis 3.

4.2.3 Hypothesis 5

Gap in the Literature: The ability of *SMCTests* to predict real-world negative driving outcomes is unknown.

Hypothesis: Poorer performance on *SMCTests* measures are associated with increased crashes and traffic offences over the following years.

Rationale: As with the rationale for Hypothesis 4, decrements in cognitive and sensory-motor performance may be predictive of future negative on-road outcomes even before the presence of a cognitive disorder is apparent.

Significance: As with Hypothesis 2, early detection of potential predictors of negative driving outcome could form part of a screening assessment to determine whether a driver is in need of additional assessment.

Study design (Chapter 7): Same as for Healthy Driver Follow-Up Hypothesis 1.

4.2.4 Hypothesis 6

Gap in the Literature: Only one study has investigated the relationship between self-reported lapses, errors, and breaking of road rules and the relationship to driving in older adults (Parker et al., 2000) and the outcome variable was self-reported crashes in the five years prior to the study. No study has looked at the relationship between these self-reported driving behaviours and prospective crashes or traffic offences.

Hypothesis: Drivers who report a higher number of driving lapses or errors on the Driver Behaviour Questionnaire over the immediate years following an on-road driving assessment will be more likely to have had a crash or traffic offence.

Rationale: As discussed in Section 3.5, self-reported driving behaviours, particularly errors and lapses, have been shown to be related to self-reported crashes in older drivers. This is in contrast to other studies which have found violations to be the biggest predictor of crashes in younger drivers. As Parker et al. (2000) suggest, this could be an indication of changes in cognition which make lapses of attention more of a problem for older drivers.

Significance: A questionnaire that aids in the prediction of on-road driving outcomes would be useful for inclusion in screening tests for older adults. Finding an association between errors or lapses of attention and the incidence of crashes and offences could also replicate

Parker et al.'s findings and suggest that ageing leads to the perpetration of errors for differing reasons than it does in younger drivers. This could reinforce the need for specific driving assessment of older drivers that are not based on detection of the violation of road rules but, rather, lapses of attention which can lead to crashes.

Study design (Chapter 7): Same as for Healthy Driver Follow-Up Hypothesis 3.

4.2.5 Hypothesis 7

Gap in the Literature: Few studies have investigated the impact of medical illness on driving, largely due to the difficulties imposed by individual responses to illness and the unpredictable effects of comorbidity.

Hypothesis: Drivers who report experiencing more distress associated with medical conditions are more likely to have crashes or traffic offences over the following years.

Rationale: Molnar et al. (2007) found that older drivers who reported being bothered a great deal by diabetes were more likely to have a crash history. This suggests that subjective distress caused by medical conditions may be a useful predictor of driving ability, perhaps beyond the simple presence of illness itself. Due to the complex nature of how illness and comorbidity affect individuals, a measure of subjective distress may work well as a measure of the functional impact of illness.

Significance: There is great difficulty in determining the extent to which medical conditions affect driving aside from some main effects found for certain conditions such as the presence of heart disease linked with driving problems (McGwin et al., 2000; Sagberg, 2006). With little evidence available to suggest the impact of medical illness, and none to guide decision-making around the effects of comorbidity, an inquiry into the stress related to medical conditions would provide a useful way to determine the increased risk of an older driver.

Study design (Chapter 7): Same as for Healthy Driver Follow-Up Hypothesis 3.

4.3 Dementia and Driving Study

The Dementia and Driving study (Chapter 8) aimed to find off-road predictors of on-road assessment Pass and Fail outcomes in a group of drivers referred to the Driving and Vehicle Assessment Service at Burwood Hospital due to diagnoses or suspected diagnoses of dementia, cognitive impairment, or memory problems. Measures included an extended cognitive test battery, questions about medical conditions and sensory-motor testing. The classification model formed was then to be tested using leave-one-out cross-validation to determine the stability of the model and estimate how accurately it would generalize to a new sample. In addition to these major goals of the study there were several specific hypotheses investigated.

4.3.1 Hypothesis 8

Gap in the Literature: There is no current study into the ability of *SMCTests* in combination with standard cognitive tests to aid in the prediction of on-road driving assessment outcome in drivers with mild cognitive impairment or Alzheimer's dementia. There are also no other computerized tests of driving ability that encompass the breadth of sensory-motor testing offered by *SMCTests*.

Hypothesis: A combination of standard cognitive tests and *SMCTests* measures will provide on-road Pass and Fail accuracy statistics higher than that achieved by previous studies.

Rationale: Mild cognitive impairment and dementia are identified by progressive deterioration in cognitive abilities as discussed in Section 2.3. As part of the normal ageing process a number of sensory and motor skills decline (see Section 2.2) which could have implications for driving ability. While cognitively-healthy older drivers may be able to adapt their driving to the changes associated with normal ageing, it is less likely that a cognitively-impaired driver will be able to do so. Therefore we may expect to see that sensory-motor measures, such as those measures using *SMCTests*, are useful additions to a predictive model of on-road driving assessment performance.

Significance: All drivers with dementia will need to stop driving at some point and there is no general consensus regarding which cognitive or sensory-motor processes are the most important to aid in determining when cessation should occur. With the expected increase in the number of adults with cognitive impairment in the coming decades, off-road predictors of driving ability will become increasingly important.

Study design (Chapter 8): 60 participants will be recruited from consecutive referrals to the Driving and Vehicle Assessment Service at Burwood Hospital, Christchurch. All will have been referred with diagnosed or suspected Alzheimer's dementia, mild cognitive impairment, nonspecific cognitive impairment, or memory problems. All will complete standard off-road testing with *SMCTests* and an on-road driving assessment. A subgroup of these drivers will complete additional extensive cognitive testing and provide additional information about their driving behaviour. This subgroup of participants will be diagnosed as MCI or Alzheimer's dementia on the basis of cognitive testing and reports of impairment in activities of daily living by family members or close friends. Measures will be used to construct models of Pass and Fail outcome on the on-road driving assessment.

4.3.2 Hypothesis 9

Gap in the Literature: No study has investigated the impact of stress related to medical illness on a group of driving with mild cognitive impairment and dementia.

Hypothesis: Drivers with MCI or Alzheimer's dementia who report experiencing more distress from medical conditions will be more likely to receive a fail score in the on-road driving assessment.

Rationale: Both MCI and dementia are neurological disorders, making any additional medical illness a candidate for the interactive negative affects of comorbidity. Distress related to other medical illnesses may impact the driving of people with MCI and dementia who are already at increased risk of experiencing problems with driving.

Significance: As stated in Hypothesis 8, the growing number of older drivers with cognitive disorders will put extra burden on the health system in deciding which drivers are likely to be unsafe drivers. Finding a relationship between bother caused by medical illness and unsafe driving in this group could be a useful predictor that could be incorporated into decisions regarding driving assessment and cessation for this patient group.

Study design (Chapter 8): Same as for Dementia and Driving Hypothesis 8.

4.3.3 Hypothesis 10

Gap in the Literature: In New Zealand there is no standardized on-road assessment for determining whether a driver with cognitive impairment is safe to drive. This raises the question of whether assessments are reliable both within raters and between raters, and whether there is a particular pattern of on-road errors that can aid in discriminating between on-road Pass and Fail groups.

Hypothesis: An on-road error list will contain errors which can discriminate between drivers with dementia or mild cognitive impairment who Pass and Fail an on-road driving assessment.

Rationale: Although MCI and dementia affect individuals in an unpredictable manner, there may be broad cognitive deficits that increase the chance of certain on-road driving errors. By examining the documented driving errors of a group of drivers with MCI and dementia it may be possible to find those errors which best discriminate between Pass and Fail groups.

Significance: Finding a list of discriminative errors could be of great use in the scoring of on-road driving assessments for drivers with cognitive impairment. Since most on-road assessors do not used standardized scoring procedures (Korner-Bitensky et al., 2006), a short list of errors could allow for a more reliable assessment of driving in those with cognitive impairment, both within assessors and between assessors.

Study design (Chapter 8): Same as for Dementia and Driving Hypothesis 8. In addition, an on-road error list will be completed by both driving assessors on completion of each on-road driving assessment.

4.4 Explanation of Hypotheses

The Healthy Driver Follow-up study and the Dementia and Driving study were largely designed following completion of the Healthy Older Drivers study. These studies benefitted from the findings of the first study and their hypotheses reflect this. For example, both the Older Drivers Follow-up study and the Dementia and Driving study include hypotheses regarding the effect of distress associated with medical conditions and driving. Although medical condition information was collected in the Healthy Older Drivers study, it was not done so in a systematic manner that allowed for anything but descriptive reporting, hence there is no hypothesis regarding medical conditions for this study. The Dementia and Driving study includes a hypothesis regarding differences in on-road errors performed between on-road Pass and Fail groups. The driving error scale used in this study (and detailed in Section 5.10.1) was compiled based on errors performed on-road during the Healthy Older Drivers study, hence, this error list was not used in the Healthy Older Drivers study.

CHAPTER 5 -
Assessment Methods

5.1 Objectives

The objective of this chapter is to list and describe the measures used in this research project.

The Healthy Older Drivers study (Chapter 6) used *SMCTests* and a selection of standard cognitive tests to construct models of classification for passing or failing an on-road driving assessment in a group of 60 cognitively-unimpaired older drivers.

The Healthy Driver Follow-up (Chapter 7) followed the participants from the first study for two years with annual telephone interviews that gathered information regarding driving behaviours, medical conditions, and crashes and traffic offences.

The Dementia and Driving study (Chapter 8) used *SMCTests* to classify on-road Pass and Fail outcome in a group of 60 drivers referred for a medical driving assessment due to suspected or diagnosed dementia, cognitive impairment or memory problems. A subset of 32 of these participants completed additional extensive cognitive testing, as well as providing information about medical conditions, driving behaviour, and being tested on road code knowledge. Many tests in the Dementia and Driving study were used to aid in the diagnosis of participants mild cognitive impairment or probable Alzheimer's dementia. The tests used for this purpose and the process of determining this diagnosis is given in Section 8.1.3.

5.2 Demographic Interview

5.2.1 Healthy Older Drivers Study

An interview was used to obtain information regarding age, education, longest held occupation, years of driving, handedness, and major medical conditions in the Healthy Older Drivers Study. Participants were asked to estimate how many km they drove a year (less than 5,000, 5,000 – 10,000, 10,000 – 14,999, 15,000 – 19,999, and 20,000 and over) and were also asked about longer car trips they had taken in the previous year. Total km travelled was estimated using tables of travel distances and Google Maps (http://maps.google.co.nz/). Km totals for additional travel were added to extrapolated driving log data obtained by multiplying the km recorded in the weekly driving log (see Section 5.3.1) by 52 weeks to form an estimate of km travelled over the previous 12 months.

5.2.2 Dementia and Driving Study

The interview was shorter in the Dementia and Driving study, due to the expected increased difficulty MCI and Alzheimer's dementia participants would have in recalling information. Participants were not asked to estimate their yearly mileage and were not asked to recall longer car trips taken in the previous 12 months. The presence of current medical conditions was investigated using a checklist (McGwin et al., 2000). Participants were asked whether they had been diagnosed with or treated in the past twelve months for any of the following: heart disease, cancer, stroke, Parkinson's disease, dementia, high blood pressure, high cholesterol, thyroid problems, sleep apnoea, diabetes, multiple sclerosis, arthritis, osteoporosis, cataracts, macular degeneration, diabetic retinopathy, glaucoma and retinal detachment; mental health conditions: depression or anxiety disorders; and a number of injuries: broken bones, surgery, falls, and head injury. If any of these medical issues were evident in the last year, they were asked two further questions: whether they were currently taking an associated medication (yes or no), and how bothered they were by their condition on a daily basis (not at all, a little, or a great deal).

5.3 Driving Questionnaires

Six driving related questionnaires were used to explore whether self-reported aspects of driving or driving knowledge were related to subsequent Pass or Fail score on the on-road driving assessment.

5.3.1 Driving Log (Healthy Older Drivers Study)

Participants were asked to log of their driving behaviour over a one week period prior to the first assessment appointment. Participants were provided with a seven-day driving log where they recorded their odometer readings before and after each driving trip. These data were used to generate a measure of driving exposure.

5.3.2 Driving Anger Scale (Healthy Older Drivers Study)

The Driving Anger Scale (DAS, Deffenbacher et al., 1994) was developed to assess the propensity for people to become angry in driving situations. In college students, higher scores on the 14-item Driving Anger Scale have been associated with increased self-reported risky driving behaviour (Deffenbacher et al., 2002; Deffenbacher et al., 2003; Dahlen & White, 2006; Schwebel et al., 2006). The original scale was developed using cluster analysis with the 14-item scale including questions for hostile gestures, illegal driving, police presence, slow driving, discourtesy, and traffic obstructions (see Appendix F for a copy of the scale). Participants rated how much anger they would feel in each situation using a five-point scale (1 = none at all, 2 = a little, 3 = some, 4 = much, 5 = very much). Since the DAS had not been used before in an older adult sample, it was completed twice by each participant: once at home before the first appointment and again during the off-road testing appointment in order to determine test-retest reliability of the measure.

5.3.3 Driver Behaviour Questionnaire (Healthy Driver Follow-Up Study)

The 24-item Driver Behaviour Questionnaire (DBQ, Reason et al., 1990) was developed to assess the frequency of occurrence of 24 self-reported driving behaviours divided into three subscales measuring areas of lapses of attention, errors, and deliberate violations of road rules (see Appendix G for a copy of the questionnaire). For each question participants were asked how often they had performed a certain driving manoeuvre in the last year (e.g., "how often have you gotten into the wrong lane when approaching a roundabout or intersection?"). Answers were rated using a six-point scale (0 = never, 1 = hardly ever, 2 = occasionally, 3 = quite often, 4 = frequently, 5 = nearly all the time).

5.3.4 Driving Habits Questionnaire (Healthy Driver Follow-Up and Dementia and Driving Studies)

5.3.4.1 Healthy Driver Follow-up Study

The Driving Habits Questionnaire (DHQ, Owsley et al., 1999) was developed to collect a range of information about a person's driving in the previous twelve months. The DHQ was modified in a number of ways from the original for The Healthy Driver Follow-up study (see Appendix H for the version used for this study). The *Exposure* section was removed as it was believed that information collected through the driving log (see Section 5.3.1) and additional questions regarding longer car journeys would provide a more accurate estimate of driving exposure. The *Dependence* section was removed as we considered the already present question "Which way do you prefer to get around?" would provide adequate information on how a participant preferred to travel (e.g., to drive themselves or be driven by others). As the questionnaire was developed in North America, references to left-hand turns were changed to right-hand turns. The *Driving Space* section was altered by removing the sixth question ("During the past year, have you driven to places outside the southeast region of the USA?"), and changing the fifth question to "During the past year, have you driven to places outside of the South Island (*of New Zealand*)?". This question included any driving performed overseas.

A number of additional questions were added to the DHQ to assess aspects of driving not covered in the original questionnaire. These included listing the forms of transport a participant used regularly (drives own car, rides as a passenger, uses a taxi, uses a motor scooter/motorcycle, uses a bicycle, takes the bus, walks places), whether the participant thought they were driving the same amount, more, or less than the previous year, whether the participant had taken driving lessons in the last year, and whether the participant's doctor had asked them any questions regarding their driving in the last year. Participants were also asked whether they experienced the same list of medical conditions described in Section 5.2.2 and were asked to self-report involvement in crashes and receipt of traffic tickets in the preceding 12 months.

5.3.4.2 Dementia and Driving Study

Participants in The Dementia and Driving study completed a pared-down version of the Driving Habits Questionnaire used in The Healthy Older Drivers study. The DHQ was shortened in order to reduce administration time for participants, and also due to problems expected with requiring people with Alzheimer's dementia and MCI to provide accurate self-reported information. From the *Current Driving* section only questions 4, 5 and 6 were asked (Do you wear glasses or contact lenses when you drive? Do you wear a seatbelt when you drive? How would you rate the quality of your driving?). The self-reported *Crashes and Citations* and *Driving Space* sections were removed due to expected inaccuracies in reporting in a sample with memory and other cognitive problems. Participants were asked about forms of transport they were using regularly and whether they experienced the list of medical conditions described in Section 5.2.2. Often an informant was present during the completion of the DHQ and was able to provide clarification or information left out by the participant.

Since the DHQ was altered substantially for both the Healthy Driver Follow-up and Dementia and Driving studies, its results should not be considered a test of the DHQ's validity for use with the participants representing these populations. Also, composite scores were not computed for some of the sections as described by the authors (Owsley et al., 1999).

The DHQ was chosen due to the inclusion of specific sections considered to be relevant to the current studies which otherwise would have had to be constructed by the primary researcher.

5.4 Driving Knowledge Tests

4.4.1 Road Sign Test (Healthy Older Drivers Study)

The Road Sign Test is printed in the document *Medical Aspects of Fitness to Drive* (New Zealand Transport Agency, 2009). The Road Sign Test is accompanied with the instructions: "If you suspect a person may be showing signs of forgetfulness or memory loss, give them this simple test on common traffic signs. A person who has trouble with this test or takes a long time to answer may need further assessment." (p. 135). A patient is shown pictures of a series of six road signs and asked what the sign means and what action a driver should take on seeing such a sign (see Figure 5-1 for adapted test stimuli).

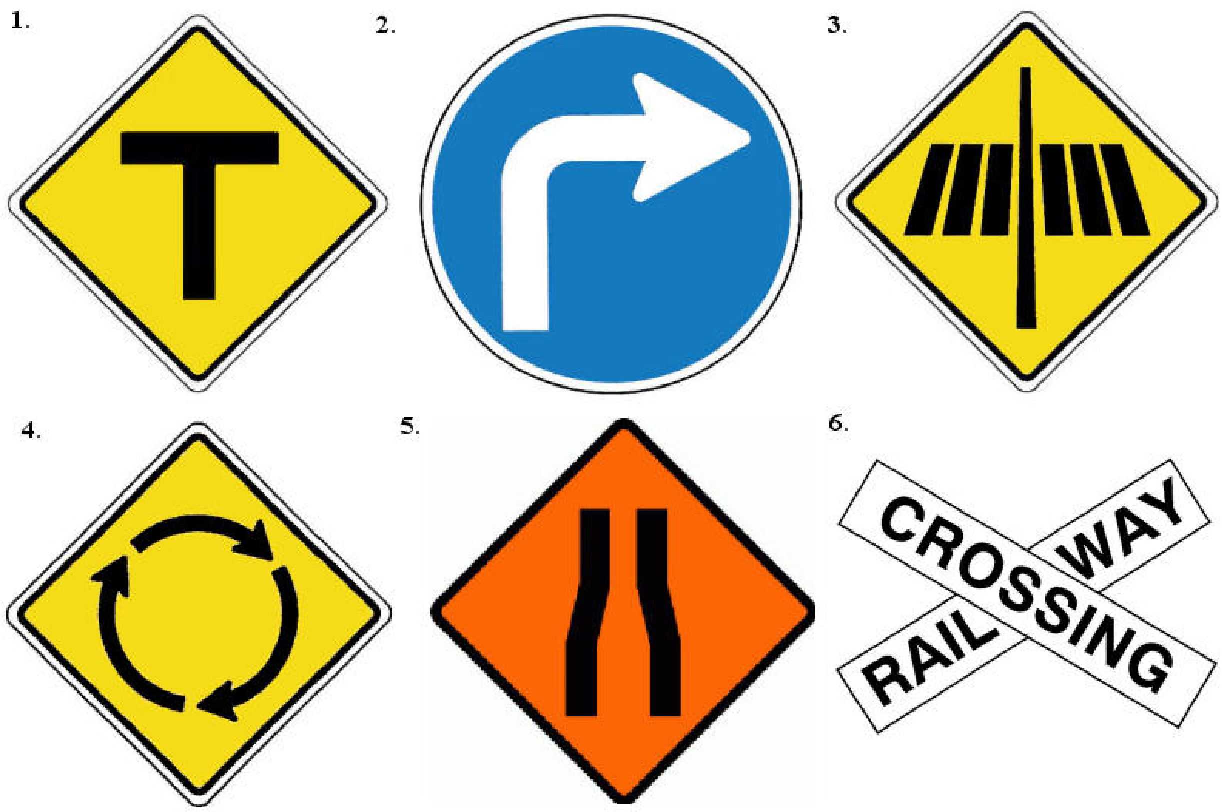


Figure 5-1 Adapted Road Sign Test stimuli in from *Medical Aspects of Fitness to Drive* (New Zealand Transport Agency, 2009) as used in the Healthy Older Drivers study.

Appropriate responses are provided but no cut-off score is offered for determining whether a patient has experienced difficulty with the test. The Road Sign Test was used in the Healthy Older Drivers study to see whether scores on the test were related to Pass or Fail outcome on the on-road assessment.

The test was presented to participants in The Healthy Older Drivers study using full colour stimuli of each sign presented individually on an A4 size sheet of paper. Each item was scored a maximum of two points (one point for correct identification of the sign and one point for correctly stating what action the driver should take) and a summed total score calculated.

4.4.2 Road Rules Test (Dementia and Driving Study)

Under New Zealand's graduated driver licensing system, learner drivers are required to complete a road rules theory test consisting of 35 random questions from the New Zealand Road Code consisting of 25 questions about road rules and road hazards and 10 questions about safety practices that relate to the type of vehicle licence the person is applying for. Thirty-two out of 35 questions (91.4%) must be answered correctly in order to pass the test. There is no additional assessment of road rules knowledge for drivers once they have achieved their learner licence. During the on-road driving assessments for The Healthy Older Drivers study it became clear from on-road driving assessment reports that many people were committing road rule violations. A selection of 15 questions from the road code test was chosen to form a brief road code knowledge test (see Appendix I). We were interested to see whether the number of road code questions correctly answered would relate to whether people passed or failed their on-road assessment in The Dementia and Driving study. Having data on how a cohort of people with dementia would fare on the road code questions could be useful for any future planning of changes to the licensing system.

5.5 Visual Acuity (Healthy Older Drivers Study)

The visual acuity of participants in the Healthy Older Drivers study was assessed using a half-size Snellen eye chart from a distance of 3 metres. Participants read the chart with each eye individually and then with both eyes together. Scoring was recorded in metric with the number recorded for analysis being the denominator in the fraction, e.g., 6/6 acuity would be recorded as '6' and is equivalent to 6/6 vision at 6 metres and 20/20 vision at 20 feet.

5.6 Sensory-Motor and Cognitive Tests (*SMCTests*™) (Healthy Older Drivers and Dementia and Driving Studies)

5.6.1 Apparatus

Two different sets of apparatus were used to complete the *SMCTests* assessment for the Healthy Older Drivers and Dementia and Driving studies. The Healthy Older Drivers study was performed on a portable system named the Canterbury Driving Assessment Tool (*CanDAT*™). Participants were seated on a non-swivel chair in front of a desk upon which sat a 19" inch LCD computer monitor for stimulus display, and under which sat a set of foot pedals (accelerator, clutch and brake). Attached to the front of the desk was a steering wheel which included turning signal indicators (see Figure 5-2). A separate laptop was used to run the *SMCTests* software which displayed test stimuli on the participant's screen and allowed the assessor to enter biographical details, administer tests, store/retrieve raw data, and analyze test performance.

Figure 5-2. The *CanDAT* hardware running *SMCTests* software used for The Healthy Older Drivers study.

For The Dementia and Driving study the *SMCTests* software was administered using a modified car body apparatus (Figure 5-3). In this device the steering wheel, indicator stick, and foot pedals (accelerator and brake) were interfaced to the A/D board of a Pentium PC. Using the PC's first graphics board, a data projector displayed 1024 x 768 pixel (80 x 60 cm) images of the test stimuli on to a plain wall directly in front of the participant with an eye-to-screen distance of approximately 180 cm. A monitor, connected to the second graphics board on the PC, allowed the assessor to run the *SMCTests* software as described for The Healthy Older Drivers study above.

Figure 5-3. The modified car body hardware running *SMCTests* software used for The Dementia and Driving study. During testing room lights are switched off and test stimuli are projected onto the wall in front of the apparatus.

5.6.2 Tests

A subset from the *SMCTests* battery was used during the Healthy Older Drivers and Dementia and Driving studies. The Healthy Older Drivers study used three sensory-motor tests (Footbrake and Clutch, Ballistic Movement, Sine and Random Tracking), and five cognitive tests (Arrows Perception, Divided Attention, Visual Search, Complex Attention and Planning). For the Dementia and Driving study, the Footbrake and Clutch and Visual Search were omitted from this list. The tests are described below and in more detail in the User Manual (*Christchurch Neurotechnology Research Programme*, 2006).

5.6.2.1 Sensory-motor function tests

Footbrake and Clutch – This test assessed a participant's ability to respond quickly to a visual stimulus. Participants were presented with a traffic light display with green and red lights. Participants were required to depress the accelerator pedal to illuminate the green light. After a random interval (2–6 s) the green light turned off concurrently with the red light turning on. Participants must then respond by lifting their foot off the accelerator and depressing the brake and clutch pedals simultaneously. Timing of the movement requires that both brake and clutch pedals be fully depressed, i.e., timing is not complete until both the accelerator and clutch pedal have reached their full depression, thus requiring strength in both legs. There were ten trials. The lowest two reaction time and movement time scores were automatically excluded in order to control for outliers, and the remaining eight reaction and movement times were averaged to give mean reaction and movement times. Times are recorded in ms.

Ballistic Movement – This test assessed reaction time, movement time, and peak velocity for arm movements in response to a non-target stimulus. Participants were required to move an on-screen arrow out of a box and across a line in response to a random 3–7 s latency stimulus (the line colour changes from red to green which was the signal for the participant to move the arrow across the line as fast as possible). Reaction time (time to respond to the stimulus change in ms) and movement time (time to cross the line stimulus in ms) were recorded in each arm separately. There were 16 trials, four for each arm in both left and right directions: the participant moved the steering wheel to the right with right arm, then left with left arm, then right with left arm and right with left arm. Mean reaction time, movement time, and total time scores were calculated using times from the 16 trials. Speed of movement wsa also recorded in mm/s. Placement of participants' hands on the steering wheel was standardized, with each participant directed to centre the arrow in the centre of the screen and grasp the top centre of the steering wheel.

Sine Tracking – This test assessed participants' ability to keep an arrow point on a sinusoidal wave which descended from the top of the screen with an 8.0 s preview before reaching the arrow point. The vertical position of the arrow on the screen was fixed, with participants able to move it left and right to track the curve of the descending line. Participants were asked to use both hands at 'ten o'clock' and 'two o'clock' positions on the steering wheel. The task required smooth movements over a 180° range of the steering wheel. The average absolute horizontal distances of the point of the arrow to the line were sampled at 60 Hz over the 70 s duration of the test and recorded in mm.

Random Tracking – This test is similar to Sine Tracking but the target is a randomly generated waveform. Again, participants were asked to keep an arrow point as close to the line as they could, with error averaged in mm over the 70 s test duration (in Figure 5-4 the random line stimulus can be seen in the centre of the screen). The Sine and Random Tracking tests were performed twice each and alternated with one another: Sine Tracking trial 1, Random Tracking trial 1, Sine Tracking trial 2, Random Tracking trial 2.

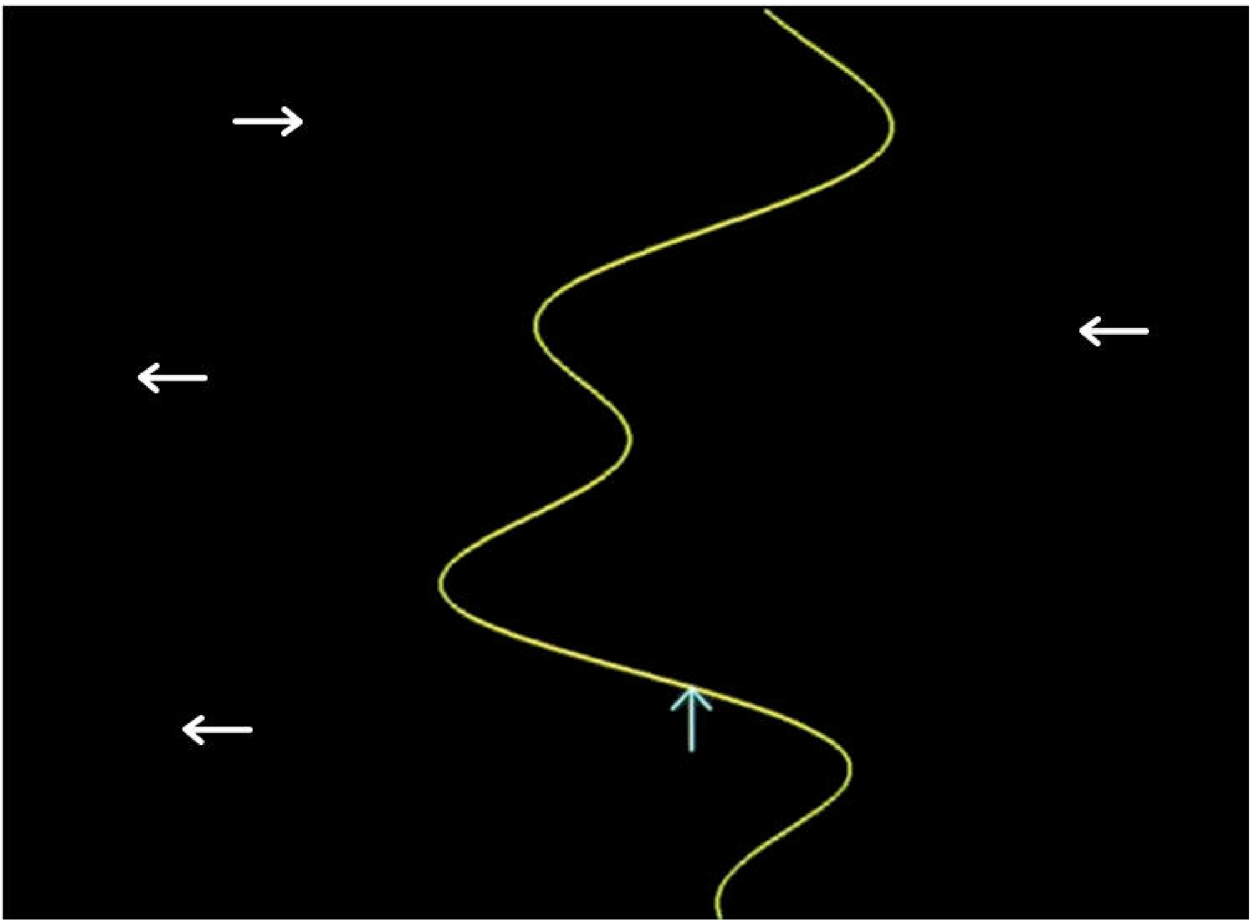


Figure 5-4. Example of the *Divided Attention* test. The participant has to follow the yellow random line target with their vertical arrow using the steering wheel while also verbally indicating whether the four horizontal arrows are pointing in the 'same' or 'different' directions. In this example the arrows are pointing in different directions.

5.6.2.2 Higher cognitive function tests

Arrows Perception – This test required participants to decide whether four simultaneously presented arrows were pointing in the same direction (all pointing left or all pointing right) or whether one or more arrows were pointing in different directions (e.g., one arrow left and three right, two arrows left and two right, etc.). The participant was directed to say “Same” if all arrows were pointing in the same direction and “Different” if one or more arrows were pointing in different directions (in Figure 5-4 the arrow stimuli can be seen at the edges of the screen). The assessor recorded the same or different decision by pressing ‘s’ or ‘d’ on the keyboard respectively. The computer recorded the time in seconds for each response, and whether the response was correct or incorrect. If the participant did not respond within 4.8 s a ‘no response’ was recorded for that trial. Each set of four arrows was displayed for a maximum of 4.8 s (otherwise a ‘no response’ trial was recorded), with a 1.0 s delay between each set. There were twelve trials in the test.

Divided Attention – This test combines simultaneous Arrows Perception and Random Tracking tests (see Figure 5-4).

Visual Search –Visual search, including left-right or central-peripheral vision bias was examined using 20 trials (see Figure 5-5). A box at the top of the screen contained the target stimuli of a left turn and right turn arrow. The participant was instructed to search for *either* of these arrows in the maintain stimulus area below. The steering wheel was to be turned in the direction that the arrow was pointing. For example, Figure 5-5 the target arrow in the array of symbols is a left turn arrow (located just above the horizontal midline on the right side of the screen). Therefore, the participant should turn the steering wheel left. Each screen was displayed for a maximum of 10.0 s.

There was an interval of approximately 2 s between each trial in which time the participant was asked to again centre their steering wheel. For each trial the response time was recorded as well as whether each steering wheel turn was in the correct or incorrect direction or whether it was not moved at all (designated a non-response trial). The mean average response time (minus any non-response trials) was reported.

Complex Attention – This test assessed participants’ ability to maintain attention on a task despite visual distracters. The test began with an arrow at the bottom of the screen (see Figure 5-6). A box at the top right of the screen contained a green light and below this was a smaller grey box at the same level as the arrow. At the top of the left side of the screen was an empty box identical to the top right hand box but without a green light stimulus. Below and to the right of the empty box was a solid green line.

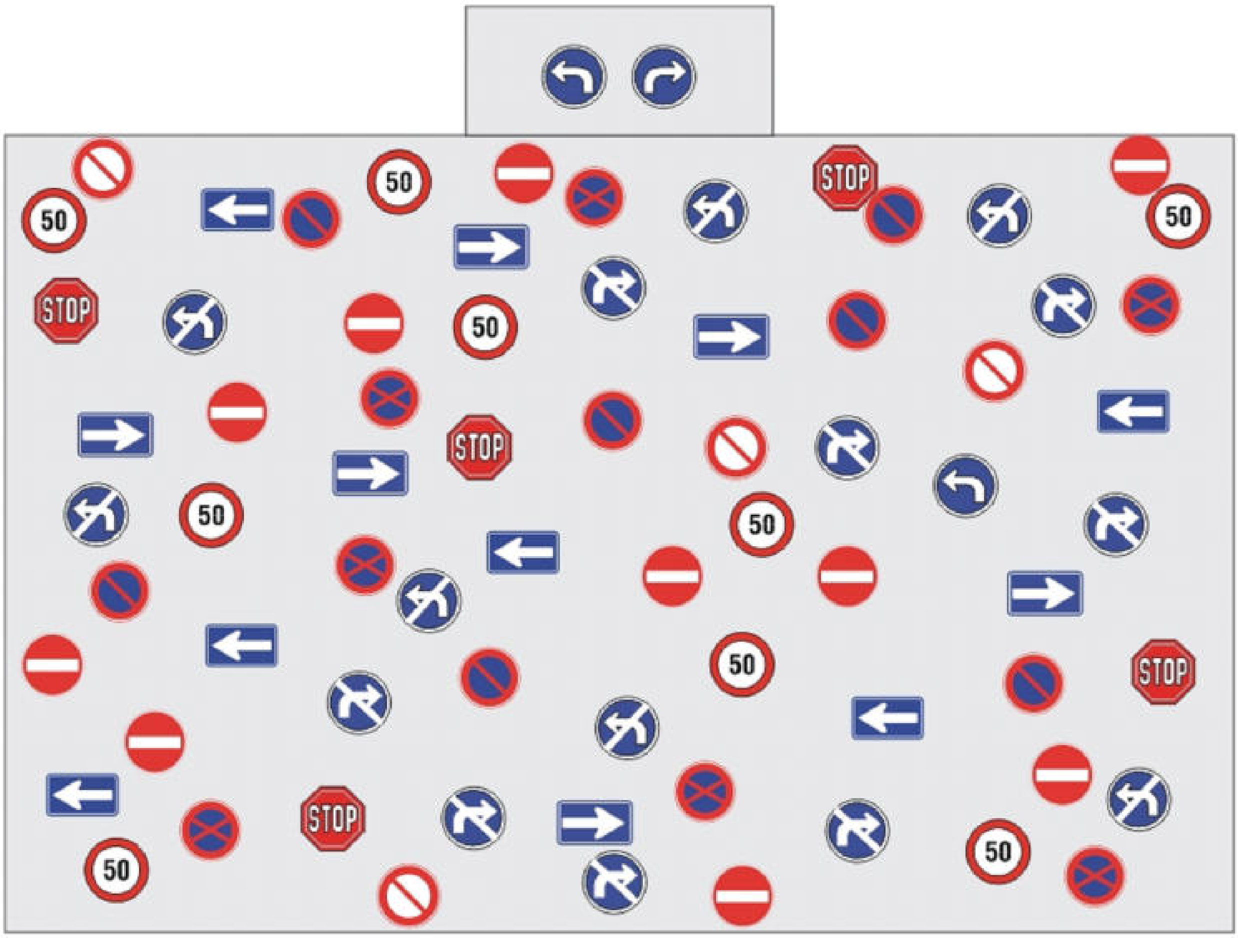


Figure 5-5. Example screen from the Visual Search test. The target stimuli are shown in the top box. A 'turn left arrow' is shown in the middle right side of the main box amongst 69 distracter stimuli.

Participants were asked to move the arrow into the small box on the right side of the screen and to remain under the green light until it moved to the grey box on the other side of the screen. At this point participants were asked to move their arrow until it was again under the green light. The top green light changed sides after a duration of between 3 and 5 s regardless of where the participant's arrow was located. The computer recorded the time from the changing light stimulus to when the participant moved the arrow out of the box as reaction time, and the time it took to move over the green light as movement time. If the participant was not within the small box at the time that the light changes or within 180 ms of the change the trial is recorded as 'invalid', and if the participant failed to move the arrow past the green line within 3.0 s of the changed stimulus the trial was recorded as a 'lapse'. The test consisted of 21 trials that took 2 min 27 s to complete.

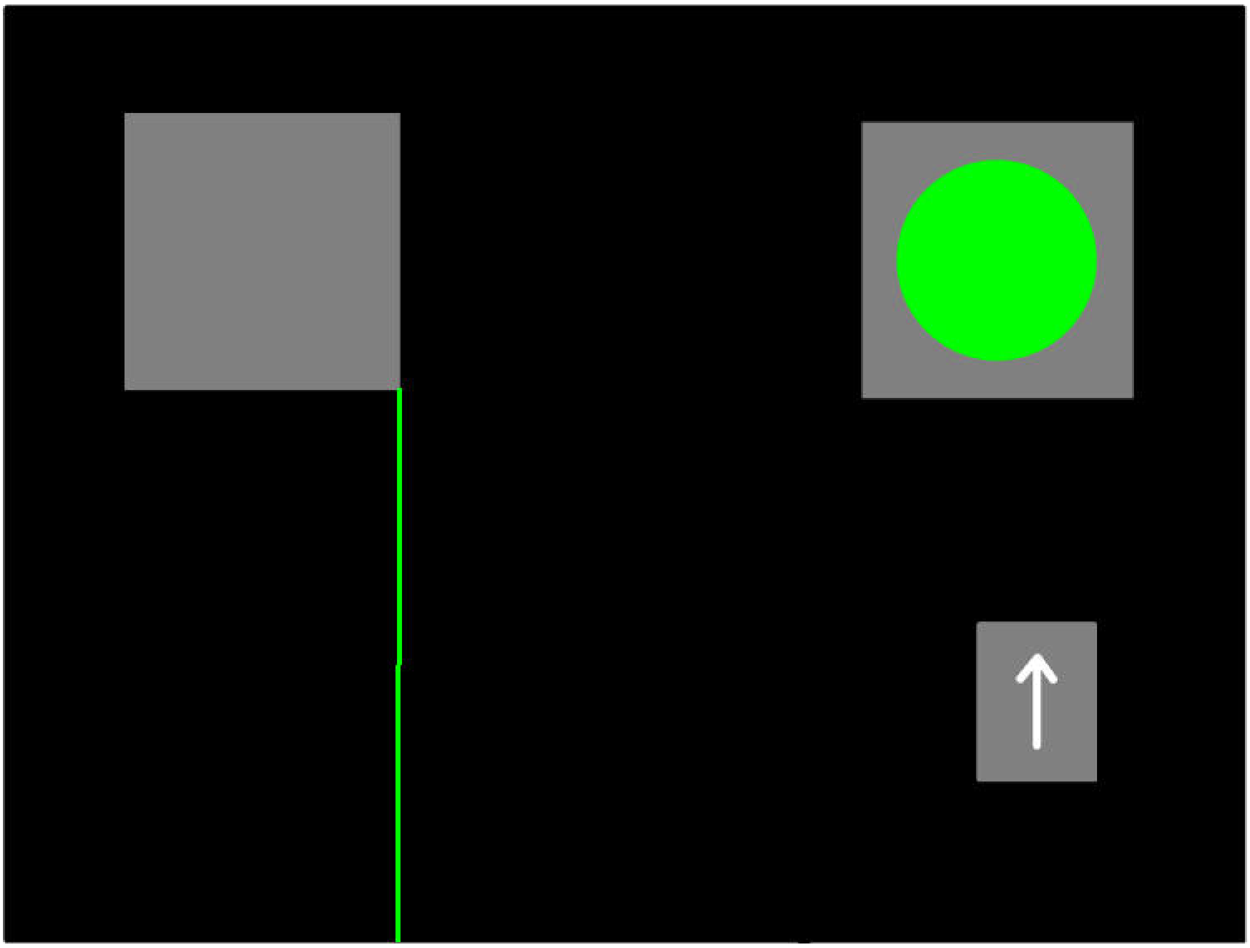


Figure 5-6. Example of on-screen presentation of Complex Attention. The participant is asked to keep their arrow under the green light.

Scores were tabulated to provide the mean reaction and movement times across trials as well as the standard deviation measure of each. The mean total time and the number of lapses and invalid trials were also recorded.

Planning – This test assessed the ability to use accurate timing and judgment to complete multi-step behavioural tasks. Participants were presented with a street scene in plan view, that consisted of a two-lane road boarded by empty green spaces ('grass') and a small blue car situated in the left lane. Participants were instructed to 'drive' the car on the road by pressing the accelerator. When the accelerator was depressed the road diagram moved down the screen at 29 mm/s, after a constant acceleration period of 1500 ms. The actual position of the blue car remained stationary on the screen. The road was usually straight except for one section of curved road. Participants were told that at some stage a hazard (puddle of spilt paint) or intersection would appear in the roadway ahead of their car (see Figure 5-7 for the car approaching an intersection).

Participants were instructed to ‘overtake’ the spilt paint hazard by selecting the right indicator (mounted on the right side of the steering wheel in the *CanDAT* setup for the Healthy Older Drivers study and in a single stalk on the right of the steering wheel in the modified car body setup for the Dementia and Driving study), moving onto the right side of the road, straightening the car, selecting the left indicator and moving back to the left lane once past the hazard. As there were cars coming towards them in the right-hand lane participants needed to determine at each hazard whether they could perform this manoeuver without stopping, or whether they needed to stop before the paint and wait for a gap in traffic to complete the passing manoeuvre.

A range of measures from Planning were utilized: lateral road position error (mm) where the standard deviation from the participant’s ‘x’ coordinate position was measured during driving on straight road sections without hazards; duration of position errors (s) where the length of time spent with any part of the blue car crossed the road was recorded; intersection safety margin (mm) where the mean distance from other cars while crossing intersections was recorded; number of hazards hit; number of crashes with other cars; and total distance travelled (m) during the test.

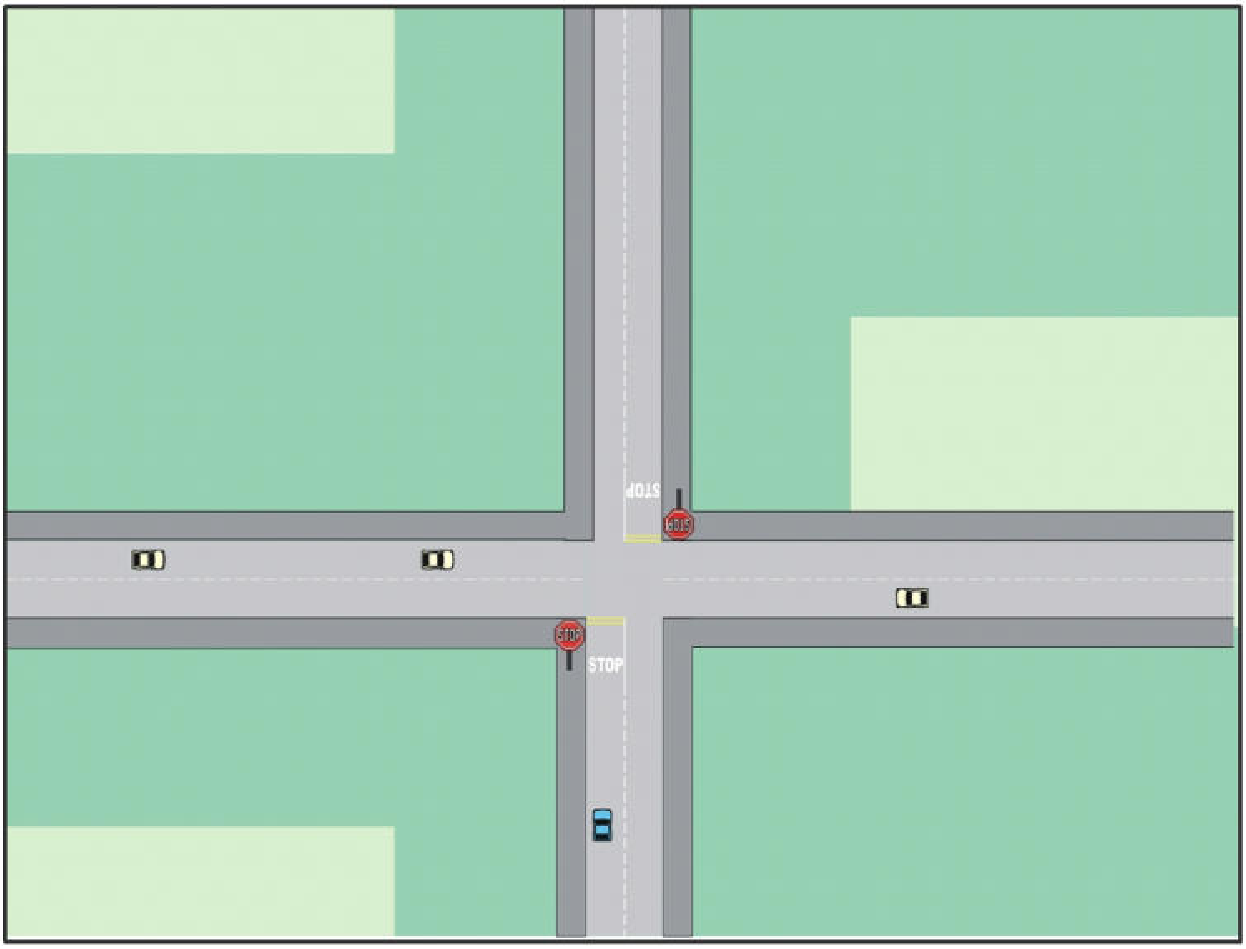


Figure 5-7. Example of on-screen presentation of Planning. In the example the participant's blue car is approaching an intersection where the participant must use the brake pedal to stop behind the yellow lines and must then choose a safe time to press the accelerator to cross the intersection.

5.7 Standard Cognitive Function Tests

A battery of standard cognitive function tests were chosen for the Healthy Older Drivers study and Dementia and Driving study. As the Healthy Older Drivers study consisted of older drivers with no diagnosed cognitive impairment, the number of cognitive tests was kept to a minimum, with those selected chosen to highlight any noticeable cognitive decline that may have been present in individuals. For the Dementia and Driving study the tests served two purposes. First, diagnoses of MCI and Alzheimer's dementia were informed by assessment results. Secondly, additional tests were added due to hypotheses about the kinds of cognitive difficulties that might be associated with the ability to drive safely.

Cognitive function tests for the Healthy Older Drivers study consisted of four tests: a standardized version of the Mini-Mental State Exam (Molloy & Standish, 1997), Trail

Making Tests A and B (*Brainmetric*, 2011), the Wechsler Test of Adult Reading (Wechsler, 2001), and the Dementia Rating Scale – 2 (Jurica et al., 2001).

Cognitive function tests for the Dementia and Driving study consisted of fourteen tests: a standardized version of the Mini-Mental State Exam (Molloy & Standish, 1997), the Wechsler Test of Adult Reading (Wechsler, 2001), two subsets of the Delis-Kaplan Executive Function System (D-KEFS) (Verbal Fluency and Colour-Word Interference) (Delis et al., 2001), Trail Making Tests A and B (*Brainmetric*, 2011), the Alzheimer's dementia Assessment Scale – Cognitive behaviour (ADAS-Cog) (Mohs, 1994), three subtests of the Visual Object and Space Perception Battery (VOSP) (Shape Detection Screening Test, Incomplete Letters, and Silhouettes) (Warrington & James, 1991), two subtests of the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III) (Letter-Number Sequencing and Block Design) (Wechsler, 1997), The Rey Complex Figure Test (copy and 3-minute recall trials) (Meyers & Meyers, 1995), and the Benton Judgement of Line Orientation test (Benton, 1983).

5.7.1 Mini-Mental State Exam (Healthy Older Drivers and Dementia and Driving Studies)

The Mini-Mental State Exam (MMSE) is a commonly used brief screen for cognitive impairment and includes 22 items covering areas of orientation to time and place (asking what year it is, the date, season of the year etc), short-term memory (asked to recall three words after a gap of a few minutes, and carry out a verbally presented command), language (asked to repeat a phrase, name two items, spell a word backwards, and write a sentence), and constructional ability (asked to copy a figure of interlocking pentagons). The maximum score is 30 with a score of below 26 often used to indicate the potential presence of cognitive impairment. The MMSE is a useful tool for detecting the presence of dementia but has poor sensitivity for detecting people with mild cognitive impairment or mild dementia who will often score in the normal range (Nasreddine et al., 2005; Smith et al., 2007). A standardized version of the Mini-Mental State Exam (SMMSE) was used, based on a version described by Molloy and Standish (Molloy & Standish, 1997), which has been shown to have higher inter-rater and test-retest reliability and is easier to administer than the original version (Folstein et al., 1975; Molloy et al., 1991).

5.7.2 *Trail Making Tests A & B (Healthy Older Drivers and Dementia and Driving Studies)*

Trail Making Tests A and B (*Brainmetric*, 2011) are used to assess visual scanning, visuomotor tracking, divided attention, and cognitive flexibility (Lezak et al., 2001). For Part A of the test, participants are asked use a pencil to link together numbers presented in circles (1 to 25) in the correct sequence (i.e., 1 connects to 2, 2 to 3, 3 to 4, etc.). Part A measures visual scanning ability which is essential in order to complete Part B of the test. Part B consists of joining together numbers and letters presented in circles in numerical and alphabetical order while switching between the two (i.e., 1 connects to A, A to 2, 2 to B, B to 3, etc.). The test was administered following instructions by Strauss et al. (2006) with time to completion (s) used as the dependent variable. In the Dementia and Driving study, those who were unable to finish the test were assigned a maximum score of 600 s. The Trail Making Test is sensitive to cognitive decline in those with dementia, traumatic brain injury, and Huntington's patients, amongst other conditions (Lezak et al., 2001). For the Dementia and Driving study the score on the Trail Making Test were converted to *z* scores based on normative data provided by Tombaugh (2004) which includes age- and education-stratified norms for a sample of non-cognitively-impaired Canadians recruited from the general population.

5.7.3 *Wechsler Test of Adult Reading (Healthy Older Drivers and Dementia and Driving Studies)*

The Wechsler Test of Adult Reading (Wechsler, 2001) consists of a list of 50 words with atypical grapheme to phoneme translations that participants are asked to read aloud (e.g., liaison, porpoise). The number of words correctly read is then converted to an estimated premorbid Full Scale IQ score. It therefore is not used in the diagnosis of dementia in The Dementia and Driving study. The ability to read words is not usually affected by injury or the early stages of dementia provided the participant had developed reading skills prior to the onset of cognitive disorder. The WTAR has been validated against the Wechsler Adult Intelligence Scale – III with correlations with Full Scale IQ of .74, .78, .72, .69, .68, and .76 with age ranges 55-64, 65-69, 70-74, 75-79, 80-84, and 85-89 respectively (Wechsler, 2001). Norms were taken from the WTAR manual using the United States 'White Male' and 'White Female' Full Scale-IQ estimates stratified by age and years of education.

5.7.4 Dementia Rating Scale – 2 (Healthy Older Drivers Study)

The Dementia Rating Scale – 2 (Jurica et al., 2001) investigates five areas of cognition that are sensitive to changes associated with dementia: attention (repeating a span of verbally presented digits, following commands, and imitation), initiation and perseveration (naming items that can be bought in a supermarket, naming article of clothing worn by the assessor, repeating back sounds, repeating movements, tapping a rhythm, and copying simple drawings), construction (copying more complex drawings, and writing their name), conceptualisation (grouping similar drawings, and naming how two items are similar), and memory (recalling a sentence after a delay, recognition of words from a word list, and recognition of designs from a design list). A total score is constructed from the results of the five subscale sections, and an age-and education-adjusted score is calculated using the Dementia Rating Scale-2 manual. A scaled score of 9 and above is described as 'intact', with scores between 6 and 8 described as mildly impaired, 4 to 5 as moderately impaired, and 3 or less as severely impaired.

5.7.5 Letter Fluency and Category Fluency (Dementia and Driving Study)

Letter and Category Fluency tests measure the rate of verbal production of words starting with three different letters (F, A, and S), and then for two different categories (names of animals and boys names). Reduced verbal fluency is often found in people with dementia (Lezak et al., 2001), and may be related to frontal lobe damage. Age-corrected norms were taken from the D-KEFS manual.

5.7.6 Colour-Word Interference (Dementia and Driving Study)

Otherwise know as the Stroop test, Colour-Word Interference measures the time it takes for participants to read a list of colour names when presented with a page containing patches of colour, to read a list of colour words presented in black ink, and to read the ink colour of words that are written in a different coloured ink from the what the word actually says (i.e., the word "red" written in green ink should be read as green). The length of time taken to read the non-congruent word and coloured ink trial and the numbers of errors made during the task are used to demonstrate difficulties with attention, concentration and/or response inhibition, which are all aspects of cognition affected by dementia. The Inhibition/Switching subtest was not used for The Dementia and Driving study. Age-corrected norms were taken from the D-KEFS manual.

5.7.7 Alzheimer's Dementia Assessment Scale – Cognitive Behaviour (Dementia and Driving Study)

The Alzheimer's dementia Assessment Scale – Cognitive (ADAS-Cog, Mohs, 1994; Mohs et al., 1997) is a more intensive screen than the MMSE or Dementia Rating Scale – 2, with 10 tests measuring word recall (10 words with 3 presentation and recall trials with a delayed recall trial a few minutes later), ability to follow four progressively more complex commands, constructional praxis (copying four shapes: circle, overlapping rectangles, rhombus and cube), naming 12 objects and five fingers, ideational praxis (folding a letter, placing it in an envelope and sealing the envelope, writing own address and showing where the stamp goes), a short list of orientation questions (date, season of the year etc), a word list followed by a recognition trial, finding a way through a short maze, and number cancellation (finding two number targets in an array of numbers). There are also four assessor-rated questions regarding difficulties in remembering test instructions, comprehending spoken language, word-finding problems, and language problems. The ADAS-Cog is most frequently used to assessment cognitive deficits over time as part of pharmaceutical trials for people with dementia. For the Dementia and Driving study subtest the total error score was used.

5.7.8 Shape Detection Screening Test, Incomplete Letters, and Silhouettes (Dementia and Driving Study)

Three subtests from the Visual Object and Space Perception Battery were used to assess deficits in object and space perception that can occur following brain damage. The Shape Detection Screening Test is used as a screen to check the participant's vision. Incomplete Letters presents a series of 20 black capital letters of which 70% of the black ink has been randomly degraded. Participants must respond as to which letter of the alphabet is being viewed. The Silhouettes test presents 15 pictures each of animals and objects which are presented in silhouette form. Silhouettes were drawn from the outlines of objects with varying degrees of angular rotation. Items are presented in order of difficulty. Norms were taken from a paper by Herrera-Guzmán et al. (2004) using healthy older adults from a Spanish population. Norms were age stratified, but were only available up to the age of 80, therefore they may underestimate the abilities of people in the 80 and above age group.

5.7.9 Letter-Number Sequencing (Dementia and Driving Study)

Letter-Number Sequencing is a subtest of the Wechsler Adult Intelligence Scale – III (WAIS-III, Wechsler, 1997) which measures working memory and attention. The task consists of the examiner reading a string of numbers and letters and asking the participant to rearrange and repeat the sequence back with numbers first in numerical order and then the letters in alphabetical order (e.g., the sequence 7-N-4-L should be repeated back as 4-7-L-N). The test starts with a list size of two and continues to a list size of seven with each list size presented with three different letter/number strings. Participants are required to fail all three strings of a certain length in order for the test to be discontinued. The score is the number of correct items. Age-corrected norms were taken from the WAIS-III manual.

5.7.10 Block Design (Dementia and Driving Study)

Block Design is a subtest of the Wechsler Adult Intelligence Scale – III (Wechsler, 1997) which measures visuospatial organisation. Participants are asked to construct two-dimensional designs using blocks with red, white, and red/white faces. Participants are first asked to copy the assessor in making a design and then construct models to match a set of drawings. Once several designs are made using four blocks, participants are asked to make designs using nine blocks. Performance on Block Design is often lowered in the presence of any kind of brain impairment. Block Design was chosen for The Dementia and Driving study due to the importance of visuospatial understanding in driving ability. Age-corrected norms were taken from the WAIS-III manual and were.

5.7.11 Rey Complex Figure Test (Dementia and Driving Study)

The Rey Complex Figure Test (Meyers & Meyers, 1995) is used as a measure of perceptual organization and visual memory. The participant is presented with a drawing of a figure comprised of both large and small details and is asked to draw a copy of the figure as well as they can. Scoring includes both the time it takes for the person to finish the drawing and the accuracy of 18 parts of the drawing each with a maximum score of 2 points (including half points), one point for accuracy of the drawing and the other for the accuracy of placement within the Complex Figure stimulus. In The Dementia and Driving study, a delay of approximately 3 minutes followed the copy trial before the participant was asked to draw as much of the figure as they could from memory. Once again, time used to complete the drawing and the accuracy and placement of 18 elements of the drawing were scored. Age-adjusted norms were taken from the Complex Figure manual.

5.7.12 Judgment of Line Orientation (Dementia and Driving Study)

The Judgement of Line Orientation test (JOLO, Benton, 1983) measures the ability of participants to estimate angular relationships between line segments. The participant is presented with eleven numbered line segments presented in a fan array and is asked to name the numbers of the two lines from the array that match two lines presented above the array (see Figure 5-8). People with dementia often perform very poorly on this test (Lezak et al., 2001), and the test was included for The Dementia and Driving study to investigate whether visuospatial deficits detected by the test are related to problems in driving. Age and sex adjusted norms were taken from Benton et al. (1994).

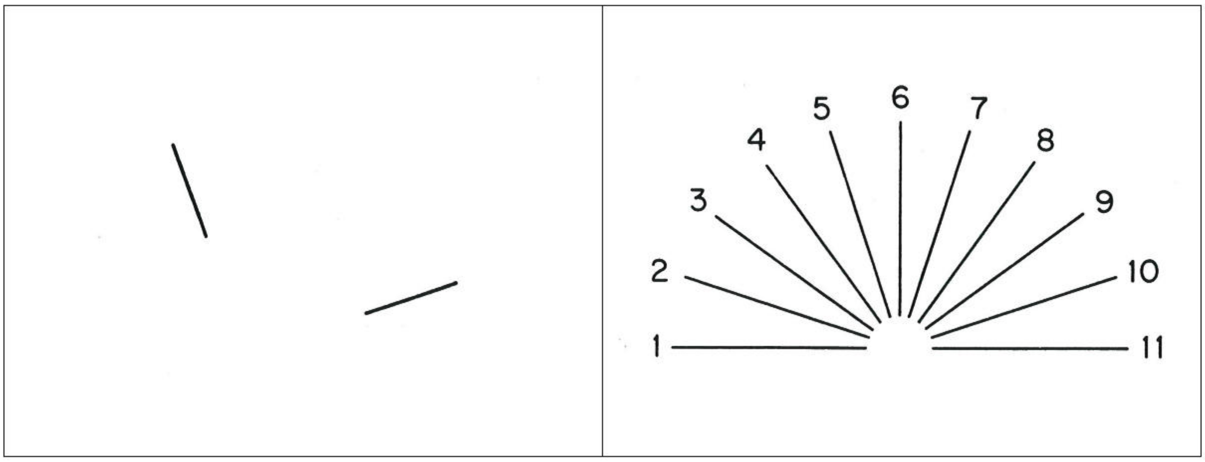


Figure 5-8. Example of a JOLO stimulus. The participant is asked to give the numbers of the two lines from the completed array (here presented on the right) that match the stimulus lines shown here on the left. The correct answer for this item are lines 5 and 10.

5.8 Psychiatric Screens and Personality Psychometrics

Apart from cognitive and sensory-motor testing, personality characteristics and symptoms of psychiatric problems may impact driving.

5.8.1 Beck Anxiety Inventory (Healthy Older Drivers Study)

The Beck Anxiety Inventory (BAI, Beck, 1990) is a measure of trait anxiety which is the propensity for people to become anxious across a range of everyday situations. In a study of young adults, trait anxiety was related to higher levels of driving anger and more risky driving and crashes (Deffenbacher et al., 2003). For the Healthy Older Drivers study, we were interested in whether those with higher levels of trait anxiety were more likely to receive a fail score on the on-road driving assessment. Raw scores of the number of items endorsed on the BAI were recorded for analysis.

5.8.2 Geriatric Depression Scale (Healthy Older Drivers and Dementia and Driving Studies)

The Geriatric Depression Scale (GDS, *Aging Clinical Research Center*, 2011) was developed specifically for older adults by removing many of the physical symptoms of depression present in other depression scales which occur at a higher frequency in a non-depressed older sample (e.g., lack of energy, problems sleeping). The screen exists in a long form of 30 questions and a short form of 15 questions. The long form was used for the Healthy Older Drivers study and the short form for the Dementia and Driving study. A previous study followed drivers aged 55 and over prospectively for five years and found that those with higher GDS scores were 2.53 times more likely to be involved in a state-recorded crash in that time compared to those with low GDS scores (Sims et al., 2000). The GDS was used as a screen to detect the presence of depression, and also to see whether there was any relationship between GDS scores and the outcome of the on-road driving assessment.

5.8.3 Big Five Personality Scale (Healthy Older Drivers Study)

The Big Five Personality Scale (John & Srivastava, 1999) was utilized to measure five personality dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience to determine whether there were any relationships between scores on these variables and on-road driving outcomes.

5.9 Activities of Daily Living Questionnaires (Dementia and Driving Study)

A diagnosis of dementia depends on the presence of significant cognitive impairment as well as impairment in activities of daily living (ADL). When cognitive impairment is present and ADLs are not significantly impaired then a diagnosis of dementia cannot be given. Instead, a diagnosis of mild cognitive impairment may be appropriate. Measures of ADLs can be divided into more simple basic tasks such as dressing, toileting and personal hygiene, and more complex activities (often called instrumental activities of daily living, IADL) which rely on higher order cognitive ability, such as memory and executive functions (Marson & Hebert, 2006). Driving is often classed as an IADL due to its use of a combination of more simple automatic procedural motor skills with intermittent demands on higher order skills such as complex attention, judgment and decision-making. For The Dementia and Driving study, it was essential to collect information about ADLs in order to discriminate between participants with MCI and Alzheimer's dementia. These measures may also be useful for determining whether a participant is likely to Pass or Fail an on-road driving assessment. Described below are the three informant questionnaires used for this task.

5.9.1 Four-Item Instrumental Activities of Daily Living Scale

The Four-Item Instrumental Activities of Daily Living scale (4IADL) was developed by Li et al. (2006) from a larger group of items assembled by Galasko et al. (1997). Li et al. (2006) administered 18 ADL questions to a family member or caregiver of participants without cognitive impairment, with diagnosed MCI or Alzheimer's dementia. Using logistic regression, the authors found a set of four items which discriminated between the two diagnostic groups. These items assess the patient's ability to find personal belongings around the house, manage their finances, keep appointments or meetings, and to read and talk about material found in books, magazines or newspapers. The addition of a single question about whether the informant thought that the patient's memory or other mental abilities had declined (answered simply 'yes' or 'no') lead to a final sensitivity for classifying MCI of 86.5% and a specificity of 79.5% and a sensitivity for classifying dementia of 85.7% and a specificity of 85.2%. Scoring is completed by adding the totals of the four standard items and subtracting the score for the single question that asks about the decline of memory and other mental abilities. Lower total scores correspond to greater impairment. The authors also provided cutoff scores for the diagnosis of MCI and Alzheimer's dementia.

5.9.2 Alzheimer's Dementia Activities of Daily Living International Scale

The Alzheimer's dementia Activities of Daily Living International Scale (ADL-IS, Reisberg et al., 2001) was developed through a process of expert review, interviews with people with dementia and their caregivers and finally a trial using controls and people with MCI and Alzheimer's dementia. The aim of the scale was to be sensitive to detecting MCI and mild Alzheimer's dementia, as well as being sensitive to change over time. The scale was developed to gather information about a large sampling of daily living activities that are affected by MCI and Alzheimer's dementia. The scale consists of 40 items from 13 different categories (conversation, recreation, self-care, house-hold activities, general activities, medication, social functioning, telephone, reading, organization, food preparation, travel, and driving) and correlates with scores on the MMSE (r = .81) and the Global Deterioration Scale

(r = .81) which measures the level of impairment for people with dementia. For each question an informant is asked about the frequency of performance of certain activities with responses rated either on a 5-point scale (0 = never, 1 = sometimes, 2 = often, 3 = always, and 4 = activity no longer performed), or as ‘activity was never performed’, or ‘unknown’. Scoring is completed by summing the scores of the items rated on the 5-point scale and dividing by the number of items rated using the 5-point scale. In this way, activities that were never performed and activities about which performance is unknown are not included in the final score. Higher scores correspond to greater impairment.

5.9.3 The Informant Questionnaire on Cognitive Decline in the Elderly

The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE, Jorm & Jacomb, 1989) facilitates assessment of cognitive decline from pre-morbid levels of functioning. The scale asks the informant to think of what their friend or relative was like 10 years ago and compare what they are like currently for a list of items such as ‘Learning new things in general’ and ‘Remembering things that have happened recently’. Each response is score on a 5-point scale (1 = much improved, 2 = a bit improved, 3 = not much change, 4 = a bit worse, and 5 = much worse). This system of scoring allows for both declines and improvements in functioning to be measured. A long form of 26 items and a short form of 16 items exist, with Jorm (2004) suggesting that the short form is most appropriate to use for the English version of the scale. Scoring consists of summing item scores and dividing by the number of item responses with higher scores indicating more impairment. A review article presents a collection of cut-points suggested by authors who have used the scale on a general population sample as well as clinical samples and suggests a cut-point of $\geq$ 3.44 has a reasonable balance of sensitivity and specificity for detecting dementia in clinical samples of around 100% and 86% respectively (Jorm, 2004).

5.10 On-Road Driving Assessment

On-road driving assessments were conducted by an experienced driving occupational therapist and a driving instructor, both from the Driving and Vehicle Assessment Service at Burwood Hospital, Christchurch. On-road assessors were blind to the results of all off-road testing. Participants were able to use their own cars (automatic or manual) for the driving assessment, as older drivers are more likely to pass an on-road driving assessment if they use their own car (Lundberg & Hakamies-Blomqvist, 2003). The driving instructor sat in the passenger seat, provided directions, and maintained safety of the vehicle while the occupational therapist sat in the rear and observed driving performance. All participants in the Healthy Older Drivers study travelled the same 45-minute public road route with an equal number of left and right turns. Participants in the Driving and Dementia study did not all drive the same route, and many started the route from their own homes (see below for more details on this). Road conditions included single-lane roundabouts, dual-lane roundabouts, dual-lane roads, controlled intersections (yield and stop signs, and traffic light controlled), uncontrolled intersections, and changes in speed zone (i.e., 50 km/hr, 60 km/hr, and 80 km/hr sections). Driving ability was rated as a consensus Pass or Fail score. The assessors were free to use any information to inform this decision. It was standard practice in driving assessment situations for occupational therapists to incorporate a range of information into their outcome decision, including a person's manner on the telephone, missing of an appointment, manner during the off- and on-road assessments, expressed concerns of family members present, and orientation to place, time, and reason for requiring a driving assessment. In the current studies the on-road assessors were blinded to all off-road testing results, and so their experience of the person's behaviour was based on their behaviour just prior to, during, and following the on-road driving assessment.

Following the Pass and Fail score decision, a driving scale score was assigned by the occupational therapist using an 11-item ordinal driving scale where scores of 0-5 could be given to those in the Fail range and scores 6-10 given to those in the Pass range (Innes et al., 2007) (see Appendix J for a copy of this scale). This scale was designed to give a continuous measure of how well a person performed in the on-road assessment.

Following the on-road assessment in the Healthy Older Drivers study, the occupational therapist provided feedback to the participant regarding driving errors and the correct behaviour for these situations. The occupational therapist was also free to recommend driving lessons for drivers if she felt there were safety concerns.

Common errors performed during the on-road assessment in the Healthy Older Drivers study were compiled into an error checklist which was then used in an effort to standardize the collection of driving error data for the Dementia and Driving study (see Section 5.10.1) In the Dementia and Driving study the on-road assessment route was not standard, with some participants beginning the assessment from Burwood Hospital and others beginning the assessment from their homes. The home option was used for people who did not have a family member able to arrange their transport, those who were considered unlikely to be able to attend an appointment due to the extent of their cognitive impairment, and those who were resistant to attending. Those who drove from home were often asked to drive a number of familiar routes rather than a novel route. Both the occupational therapist and the driving instructor monitored driving errors and filled in a driving error form following the drive.

5.10.1 On-Road Error List (Dementia and Driving Study)

An on-road error list with a standardized recording procedure was constructed based on the noted errors from on-road reports for the participants in the Healthy Older Drivers study. The form contained the 28 most frequently performed driving with extra spaces provided for assessors to write in any additional errors they thought were worth noting. The list was then given to the driving instructor in the Healthy Older Drivers study and two occupational therapists to add other errors that they commonly observed in drivers with dementia. For the final list (see Appendix K), assessors were asked to indicate errors with check marks under three categories: ‘error present’ which simply noted the performance of an error, ‘persisted despite instruction’ for errors that persisted following corrective feedback, and ‘contributed to fail judgment’ for those errors that were deemed serious enough to have directly contributed to a Fail rating. Assessors were also asked to rate the participant’s level of insight, awareness of any driving problems and receptiveness to feedback. These additional ratings were scored as ‘yes’, ‘limited’, or ‘no’. Assessors also rated whether a participant could benefit from driving lessons, as either ‘yes’ or ‘no’.

CHAPTER 6 -
Study 1 – Healthy Older Drivers

This chapter explores a range of cognitive, sensory-motor, personality, demographic and *SMCTests* measures to determine their relationship to on-road assessment Pass or Fail outcome in a group of cognitively-unimpaired drivers aged 70 and over. Following detection of the variables associated with Pass and Fail outcome a classification model was constructed using binary logistic regression which was then checked for stability using leave-one-out cross-validation.

6.1 Methods

Ethical approval was obtained for this study through the Upper South A Canterbury Ethics Committee.

6.1.1 Participants

A convenience sample of current drivers was recruited from churches, recreational groups, word of mouth, and advertisements placed in two free local health magazines in Christchurch. Participants were mailed an information sheet about the study (see Appendix L) and all participants gave informed consent (see Appendix M for a copy of the study consent form).

Sixty participants (30 males and 30 females) aged 70 to 84 years were recruited to obtain 10 men and 10 women in each of three age ranges (70-74, 75-79, and 80+ years; mean=76.7 years); 93% identified their ethnicity as New Zealand European. All were current drivers, with an average of 55.1 years of driving experience. Exclusion criteria included a history of moderate to severe brain injury, diagnosed neurological or cognitive disorder, musculoskeletal disease that interfered with driving, and recent psychiatric disorder. Other illness was not an exclusion criteria and a number of past and current health problems were reported: high blood pressure (59.3%), arthritis (55.9%), high cholesterol (39.0%), cataracts (32.2%), heart disease (27.1%), and cancer (22.0%). Males had more driving experience than females (58.8 versus 51.5 years, z = -3.82, p < .001), and drove more km per year (male median=8,693 versus 5,894, z = -2.81, p = .005). All participants scored above 26 on the SMMSE (mean=28.8), suggesting that cognitive impairment at a level indicative of dementia was unlikely. Participants were free to continue driving irrespective of the outcome of the on-road driving assessment and received NZ$50 compensation for their involvement. Table 6-1 provides a summary of participant demographics.

Table 6-1. Participant demographics by age group (total n=60)

	Age Group (years)		
	70-74	75-79	80 plus
Number in group	20	20	20
Years of age - mean (range)	71.7 (70-74)	76.7 (75-79)	81.7 (80-84)
Percent female / male	50%	50%	50%
Years of driving - mean (range)	51.9 (40-58)	53.7 (31-64)	56.5 (50-69)
Years of education - mean (range)	12.8 (8-19)	12.6 (8-19)	13.7 (9-19)
Km driven per year - median (range)	7,057 (624-29,789)	7,480 (2964-122,572)	7,142 (1128-27,034)
Ethnicity[1] - NZ European	19	19	18
- Other European	1	0	0
- Māori	1	1	3

[1]Numbers do not add to total number in each group as more than one ethnicity could be selected

6.1.2 Assessment Procedure

For the off-road assessment, three self-report measures were completed by the participant at home (Big Five Inventory, Driving Anger Scale, and the one-week driving log – see Chapter 5, Assessment Methods for all measures used in this study). At the off-road appointment participants provided information about age, years of education, longest held occupation, an estimate of km driven in the previous 12 months, reported longer car trips they had driven during the last 12 months, self-reported handedness (left or right), and major medical conditions. They also completed the Road Sign Test, Snellen eye chart, standard cognitive tests comprising the Standardized Mini-Mental State Examination, Trail Making Tests A & B, Wechsler Test of Adult Reading, Dementia Rating Scale – 2, and psychiatric screens comprising the Beck Anxiety Inventory and the Geriatric Depression Scale. The Driving Anger Scale was repeated at the first assessment appointment to determine test-retest reliability. Participants completed the *SMCTests* battery comprising the substests Footbrake and Clutch, Ballistic Movement, Sine and Random Tracking, Arrows Perception, Divided Attention, Visual Search, Complex Attention and Planning. Each participant was randomly assigned to complete either *SMCTests* or the standard cognitive measures first, to control for fatigue and order effects.

On-road assessments were conducted an average of 14.6 days (SD = 11.3, range 2-41 days) after the off-road assessment. On-road assessment were administered by an experienced driving occupational therapist and a driving instructor blinded to off-road test performance, both from DAVAS (details of the on-road assessment are described in Chapter 5, Assessment Methods).

6.1.3 Data Analysis Methods

6.1.3.1 Testing Normality of Test Data

The normality of the distribution of the data was assessed using the Shapiro-Wilk W test. Test data that was not normally distributed was examined using non-parametric statistics such as Mann-Whitney *U* tests, while normally distributed data was examined using parametric statistics such as *t* tests.

6.1.3.2 Binary Logistic Regression (BLR)

BLR is a modelling technique used to classify or predict a dichotomous dependent variable using a set of independent variables which can be continuous, ordinal, dichotomous or of normal or non-normal distribution (Hosmer & Lemeshow, 2000). In the current study, the dichotomous variable being predicted was Pass or Fail score on the on-road driving assessment.

Like all modelling techniques BLR is sensitive to overfitting, which occurs when the model fits random variance or specific relationships found within the training data which leads to reduced ability of the model to generalize to a new sample (Babyak, 2004). Babyak (2004) has several objections to the way regression models are often utilized. These will be listed, and then the design of the current study will be described.

Babyak (2004) states that choosing variables to enter into a regression model based on the relationship of independent variables to the dependent variable (he calls this ‘cherry picking’) is a post hoc strategy that greatly increases the chances that the resultant classification model will be over-fitted to training sample and may generalize poorly to a new sample. Babyak suggests that the most robust models are constructed using strong a-priori hypotheses about which variables will be useful in the regression model, and using these variables without recourse to their relationship to the independent variable in the collected data set.

Determining which independent variables are likely to be related to the dependent variable can be performed in two ways. Firstly, if there is strong a priori research base that finds specific independent variables to be useful, these can be entered into the model. Secondly, a researcher can collect data from participants and analyze the relationships between independent and dependent variables. Then a new independent group of participants should be recruited and the independent variables found to have relationships with the outcome variable in the first sample should be entered a priori into the model.

Babyak (2004) has further suggestions regarding the type of regression model variable selection process to use. He claims that using forwards and backwards selection methods further increase the risk of overfitting and suggests instead that enough participants should be recruited in order to offer variables to the model at a ratio of one variable per 10 or 15 participants. This would mean, for example, that a sample of 200 participants would allow the offering of 20 or 13 variables to the model, depending on whether the 1/10 or 1/15 ratio is used. Babyak says that models should then use the 'enter' option for these variables, meaning that all variables are forced into the model rather than letting the model choose which variables to accept as in a step-wise selection procedure. This again is supposed to reduce the risk of overfitting.

Furthermore, Babyak suggests that a form of cross-validation, such as boot-strapping, should be run to estimate the ability of the model to generalize to new data. This is supported by Steyerberg et al. (2001), who found that cross-validation models provide a better estimate of a model's ability to generalize to a new sample than split-half models which train a model on a subset of data and then test it on the held back sample, and Innes et al. (Innes et al., in press)

The suggestions of Babyak certainly go against the methods used by the majority of driving researchers, which primarily consist of collecting data from a sample, finding the independent variables most associated with the dependent variable (often by choosing those which have significant associations with the dependent variable), offering these variables to the model (without necessarily taking into consideration the ratio of participants to variables entered), and producing a classification model without using cross-validation to test its ability to generalize to a new sample. These procedures undoubtedly increase overfitting, and the lack of cross-validation likely leads to a large over-estimate of the ability of the resulting models to predict a driving outcome in a new sample.

We were either not able to, or chose not to, follow all the suggestions of Babyak for the current studies. We list the processes used in constructing the BLR models for both the Healthy Older Drivers and Dementia and Driving study (Chapter 8) below in order to make our process transparent, and to openly acknowledge the potential problems that may have arisen. Before listing the process in steps, it should be noted that we did not have the resources to recruit two independent samples, one to look at the relationship between independent variables and the dependent variable, and another to build a model from the relationships found in the first sample. Neither did we have a strong set of predictors for on-road driving ability, especially since we intended to use *SMCTests* which had not previously been used on a sample of healthy older drivers or drivers specifically diagnosed with Alzheimer's dementia or MCI. Thus, we constructed and tested our model based on the training set and thus 'cherry picked' the variables that we offered to the model. The process we undertook was as follows:

1. Variables were pragmatically considered not to be useful in offering to the model were deleted from the variable list.

This had the effect of reducing the number of variables that would contend for possible inclusion into the BLR model. These decisions were made on a number of grounds, for example, items that would never be accepted for use in a driving assessment service to use as off-road predictors of driving were excluded. This included items such as the Big Five Personality factors, and occupation code. Variables that had an effect in the opposite direction from expected were also removed, as we would not be able to adequately describe their presence if accepted into the model. This was determined by examining the means of the Pass and Fail groups for normally-distributed variables and the ranked means of the Pass and Fail groups for the non-normally distributed variables. Some variables that were subtests of a larger test were excluded in favour of the total test score as this was expected to be more likely to generalize in a new sample. All these excluded variables were still investigated in terms of their relationship to the on-road assessment outcome, and could still provide useful information about possible inclusion in future models.

2. Variables were ranked by effect size.

Since we had decided that a model would be constructed from the test data we had to have a way to 'cherry pick' the variables that were most related to the outcome. Since effect size is a more stable measure than statistical significance, which is dependent on sample size, effect size was chosen as the selection method for variables to offer to the model.

3. Variables with high collinearity with one or more other variables were deleted.

Multicollinearity found in the variables was examined using the 'Collinearity diagnostics' function in SPSS. These statistics measure the degree of collinearity among all variables entered into the equation. These relationships are independent of the relationship of variables to the dependent variable. The lower the tolerance value reported in the table, the more correlated a measure is with one or more of the other variables. Garson (2010) suggests a rule of thumb of Tolerance <0.20 for detecting variables with multicollinearity problems. Variables with the lowest tolerance values can be deleted one at a time and the analysis rerun until all independent variables have tolerance values of $>.20$.

4. Remaining variables were offered to the model in order of effect size at the ratio of 1 variable per 5 participants.

We chose a ratio of variables entered to participants of 1 to 5, which is lower than that recommended by Babyak (2004), but has been suggested as appropriate by Tabachnik and Fidell (2001). This meant that, for a sample of 60 participants, 12 variables would be selected to be offered to the model.

5. A model was formed using logistic regression with a backwards elimination procedure (criterion for entry to model $p = .05$, removal from model $p = .10$).

We decided to retain the use of step-wise selection since the aim of the study was to find a subtest of tests that could be used to predict driving and we did not want to use an 'enter' method whereby variables that did not explain significant amounts of the variance in the on-road driving assessment outcome were included. We used backwards elimination, which has been recommended over forwards step-wise selection (Steyerberg et al., 2003).

6. BLR models were tested for stability and generalizability using leave-one-out cross-validation.

The most important part of the data analysis was to estimate the ability of the classification model to generalize to a new sample. As mentioned, this is a task that is rarely completed in published driving research literature (discussed in detail in Section 6.3.5), but is essential to estimate the true predictive accuracy and usefulness of any constructed model. The method used was leave-one-out cross-validation, which is covered in detail in Section 6.1.3.5.

Babyak (2004) would likely find our method insufficient for constructing a valid and generalizable model for predicting on-road driving Pass and Fail outcomes. He would likely agree, however, with the steps of pragmatic reduction of variables, checks of multicollinearity, and cross-validation of the model. The development of any model must be met with a fair amount of caution before that model has been validated on an independent sample. The descriptive statistics and associations between the independent and dependent variables are useful for determining future studies and determining which variables provide useful information about driving ability.

6.1.3.3 Classification Versus Estimated Predictive Accuracy

One of the most important purposes of a model is to generalize the results to predict future behaviour of an independent group of people. Modelling begins by constructing an in-sample, or classification, model which is a parsimonious fit of independent variables in order to explain the score on the dependent variable. To determine whether a model is able to correctly predict outcomes for data not included in the original classification model it needs to be either tested on a new sample or investigated using statistical procedures such as boot-strapping, n-fold, or leave-one-out cross-validation. Throughout this thesis the term 'classification' is used to describe the parsimonious fitting of a model to the data it was trained on and 'predictive accuracy' or 'estimated predictive accuracy' is used to describe the process of testing the model against independent cases that it was not trained on or using statistical re-sampling methods in order to estimate its generalization to a new sample.

In the current study, the relatively small size of the study sample precluded estimating predictive accuracy using a held back sample of data, so the ability of the classification models to generalize to new data was estimated using leave-one-out cross-validation (see Section 6.1.3.5 for an explanation of this procedure). As classification models are by definition optimized to the specific characteristics of the study sample, it was expected that the estimated predictive accuracy of the BLR model would be lower than for classification.

6.1.3.4 Receiver Operating Characteristic (ROC) Curve

An ROC curve is a graphical plot of the true positive rate (sensitivity) versus the false positive rate (1-specificity) for classifying cases against a dichotomous dependent outcome measure across the range of possible cut-points for the model. An area under the curve (AUC) value of .50 indicates a model has no ability to discriminate between the outcome measure (which is Pass or Fail of the on-road assessment in the current study). A value of 1.0 indicates perfect discrimination, with values between .50 and 1.0 indicating increasing strength of discrimination. The ROC curve is plotted by entering an independent variable and a state variable. For BLR the independent variable is the predicted probability that individual participants will be in the Fail group. This data is provided as part of the logistic regression analysis.

6.1.3.5 Leave-One-Out Cross-Validation

The process of leave-one-out cross-validation consists of removing each case individually from a sample, re-training the model on the remaining participants, and testing the prediction on the excluded case using the new model (Witten & Frank, 2000). The procedure is repeated for all cases and accuracy rates averaged across all iterations. In essence, it mimics what would happen if a case was not part of the training data set and, therefore, estimates how the model would perform given a new sample from the same population (provided the sample is representative of the population). The procedure is performed as many times as there are cases in the sample.

Apart from estimating how the model would perform given a new sample, the process of leave-one-out cross-validation also provides an estimate of the stability of the model. It can do this because the procedure of leave-one-out cross-validation is for all variables offered to the full-sample classification model to be offered to all of the leave-one-out cross-validation iterations. With a sample size of 60 this means that all 12 variables (a ratio of 1 variable to 5 participants) would be offered to all 60 iterations, regardless of which variables were selected by the full-sample classification model. This procedure allows for the stability of the full-sample selected model to be estimated. For example, if a large proportion of the 60 iterations dropped measures that were included in the full-sample classification model and included other measures that were not accepted into this model we would be less certain of the stability of the constructed model as the variables included may be idiosyncratically related to particular participants within the sample.

AUC values were not calculated for models following leave-one-out cross-validation. Since the AUC describes the discriminative usefulness of a model we decided that calculating a single AUC for the culmination of different model iterations contained within a leave-one-out cross-validation analysis may not be statistically acceptable.

6.1.3.6 Choice of Cut-Points for Reporting Accuracy

Inspection of the ROC curve coordinates for each model allowed for the selection of criterion cut-points for classifying Pass and Fail outcomes which can be selected based on their overall accuracy as well as sensitivity for correctly detecting Fails and specificity for correctly detecting Passes. From the range of cut-points available, it was decided that two cut-points would be inspected for each model. The first cut-point assessed would be the default cut-point for a Fail score of 0.5 for the BLR. The second would be the cut-point that represented the highest value of sensitivity and specificity when averaged together.

We were also interested in the negative predictive value (false negatives) and positive predictive value (false positives) of each cut-point. The negative predictive value represents the proportion of all participants predicted to Pass who were actual Passes and the and the positive predictive value represents proportion of all participants predicted to Fail who were actual Fails. These values are dependent on the base rate of Passes in the population. Due to differences in reported Pass and Fail rates on on-road tests in the literature the Pass rate found in the current study will be used as the base rate for calculation of negative and positive predictive values. Ideally we would want the positive and negative predictive values of the cut-points to be as high as possible as false positive detections lead to unnecessary on-road assessments and false negative lead to Fails being classified as Passes and allowed to continue driving.

6.1.3.7 Terminology for Reporting Accuracy Statistics

Throughout this thesis the terms ‘sensitivity’, ‘specificity’, ‘accuracy’, ‘positive predictive value’, and ‘negative predictive value’ are used to present a consistent definition of the performance of the models discussed (Table 6-2).

The term ‘accuracy’ is used to denote the proportion of participants who were correctly classified as Pass and Fail on the dependent measure (i.e., on-road driving assessment). With reference to Table 6-2, this is determined by the equation ‘accuracy’ = a+d / (a+b+c+d)

Table 6-2. A representation of the values used to determine values reported as accuracy statistics

	Predicted outcome	
Observed outcome	Pass	Fail
Pass	a	b
Fail	c	d

The term ‘sensitivity’ is the ‘true positive’ rate and is used to denote the proportion of participants who failed the on-road assessment who were correctly classified as Fail on the dependent measure. With reference to Table 6-2, this is determined by the equation ‘sensitivity’ = d / (c+d).

The term ‘specificity’ is the ‘true negative’ rate and is used to denote the proportion of participants who passed the on-road assessment who were correctly classified as Pass on the dependent measure. With reference to Table 6-2, this is determined by the equation ‘specificity’ = a / (a+b).

The term ‘positive predictive value’ (also known as ‘selectivity’) is the ‘false positive’ rate and is used to determine the proportion of participants classified as Fail who actually go on to Fail on the dependent measure. With reference to Table 6-2, this is determined by the equation ‘positive predictive value’ = d / (b+d).

The term ‘negative predictive value’ is the ‘false negative’ rate and is used to determine the proportion of participants classified as Pass who actually go on to Pass on the dependent measure. With reference to Table 6-2, this is determined by the equation ‘negative predictive value’ = a / (a+c).

6.2 Results

6.2.1 On-Road Assessment

Sixteen of the 60 participants (27%) failed the on-road driving assessment with no difference in failure rates between males and females (7 males, 9 females; Fisher’s Exact Test, two-tailed p = .77). The mean age of the Fail group was 77.8 years and 76.3 years for the Pass group (Mann-Whitney U test, two-tailed, z = -1.15, p = .25). Fifty-nine participants completed the assessment in a familiar vehicle, with only one participant choosing to drive an unfamiliar car as he wanted to complete testing on a manual-transmission vehicle rather than the automatic-transmission vehicle he was currently driving. Table 6-3 summarizes the results of the Pass and Fail groups, including the number of drivers rated at each level of the driving scale score.

The driving scale scores cluster around the centre of the scale, with 60% of the sample receiving a score of either 5 or 6, just on either side of the Pass/Fail divide with a score of 7 being the second most common score after the mode score of 6.

Cronbach's α was calculated for several ordinal scales in order to determine the internal consistency of the measures. Cronbach's α increases as intercorrelations among test items increase and is generally thought to measure how well individual items in a scale are measuring a unitary construct, i.e., whether items in a depression scale all appear to be measuring a construct of depression rather than an unrelated construct.

Table 6-3. Characteristics of on-road assessment Pass and Fail groups

	Driving Scale score	Number of participants	Total	Sex	Age
Fail	0	0	n=16 Fail (26.7% of sample)	n=9 females (30% of females)	Mean age 77.81 years
	1	0			
	2	2			
	3	3		n=7 males (23.3% of males)	
	4	2			
	5	9			
Pass	6	27	n=44 Pass (73.3% of sample)	n=21 females (70% of females)	Mean age 76.25 years
	7	12			
	8	4		n=23 males (76.7% of males)	
	9	1			
	10	0			

When a scale is designed to measure a hypothesized unitary construct we would expect Cronbach's α to be approximately 0.80 or higher indicating that the sum of individual test items appear to be measuring a singular construct. The achieved Cronbach's α values were: Geriatric Depression Scale α = .79; Beck Anxiety Inventory α = .81; Driver Anger Scale first administration α = .91; Driver Anger Scale second administration α = .92; Big Five Inventory Extraversion subscale α = .80; Big Five Inventory Conscientiousness subscale α = .84; Big Five Inventory Neuroticism subscale α = .70; Big Five Inventory Openness to Experience subscale α = .73. With the exception of the Big Five Neuroticism and Openness to Experience subscales, the α values for the ordinal measures were high and therefore we can have some confidence that the items are measuring a singular construct in this older age sample.

Data from the first and second administrations of the Driving Anger Scale were correlated to investigate the form's test-retest reliability. The second administration of the scale was performed an average of 10.7 days (SD = 11.3, range 0-49) after the first. A Pearson's correlation was performed with a result of r = .80 indicating a moderately strong relationship between scores of the first and second administrations. This moderately strong test-retest reliability along with a high Cronbach's α value indicates that the Driving Anger Scale appears to be suitable for use in an older age sample.

Next, test data were tested for normality, with 78% of test measures returning significant Shapiro-Wilk W scores, indicating that the majority of the data were not normally distributed and therefore would violate the assumption of normality required for parametric testing. These results are displayed in Appendix N.

6.2.2 Pass and Fail Groups: Significant Differences and Effect Sizes

Details of the relationships between independent variables and on-road Pass and Fail score are given in Table 6-4, Table 6-5, and Table 6-6 along with Cohen's effect sizes for parametric data and Cohen's effect-size for rank-transformed variables for non-parametric data (Hopkins, 2004). A positive effect size means that a higher score was associated with an increased likelihood of failing the on-road assessment, and a negative effect size means that a lower score was associated with an increased likelihood of failing the on-road assessment. Positive and negative effect sizes do not represent whether higher or lower scores are 'better' or 'worse'. This designation must be made on the basis of each measure, e.g., a higher reaction time or error score is 'worse' while a higher peak velocity or Wechsler-type scaled score is 'better'.

Two non-normally distributed measures were significantly different between Pass and Fail groups. These were Trail Making Test B with the Fail group having longer completion times (z = -2.567, p = .010) and the *SMCTests* measure Random Tracking run 1 with the Fail group having a higher error score (z = -2.340, p = .019).

Table 6-4. Comparison of on-road Pass and Fail groups for non-normally distributed standard off-road tests using Mann-Whitney *U* tests.

Test Measure	Median for Pass group (n=44)	Median for Fail group (n=16)	Mann-Whitney *U* *p*-value **(**p*<.05)**	Cohen-type effect size[1]
Gender (1 = male, 2 = female)	1.0	2.0	.563	0.17
Age (years)	76.0	78.0	.251	0.35
Age Grouping (1 = 70-74, 2 = 75-79, 3 = 80+)	2.0	2.5	.156	0.41
Handedness (1 = right, 2 = left)	1.0	1.0	.216	-0.44
Years of education	13.0	14.0	.650	0.14
Occupation code (range 1-8 - higher = more professional occupation)	2.0	2.5	.199	0.38
Km driven last 12 months	7310	7176	.341	-0.27
Vision[2]				
Left eye	9.0	9.0	.888	0.04
Right eye	6.0	7.5	.508	0.20
Binocular	6.0	6.0	.545	0.17
Road sign test (no correct)	12.0	10.5	.432	-0.23
Mini-Mental State Exam	29.0	28.5	.701	-0.11
Geriatric Depression Scale	3.0	2.5	.637	0.12
Beck Anxiety Inventory	4.0	3.5	.403	0.25
Driving Anger Scale time 2	32.0	30.0	.200	-0.40
Trail Making Test A (s)	32.0	37.0	.055	0.58
Trail Making Test B (s)	88.5	111.5	**.010***	0.80
Wechsler Test of Adult Reading (estimated IQ score)	112.5	110.5	.610	-0.15

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value. [2]Rated using a Snellen eye chart with a score of 6 equal to a metric 6/6 vision (20/20).

Table 6-5. Comparison of on-road Pass and Fail groups for non-normally distributed *SMCTests* using Mann-Whitney *U* tests

Test Measure	Median for Pass group (n=44)	Median for Fail group (n=16)	Mann-Whitney *U* *p*-value (***p*<.05**)	Cohen-type effect size[1]
Footbrake and Clutch Test				
Mean reaction time (ms)	2965	295	.920	0.29
Mean movement time (ms)	280	304	.947	-0.02
Total reaction and movement times (ms)	578	583	.933	-0.02
Ballistic Movement Test				
Reaction time, right hand (ms)	344	354	.802	0.07
Reaction time, left hand (ms)	333	337	.616	0.15
Reaction time, grand mean (ms)	342	349	.802	0.07
Movement time, right hand (ms)	231	212	.504	-0.19
Movement time, left hand (ms)	231	223	.967	0.01
Movement time, grand mean (ms)	229	218	.593	-0.16
Total reaction and movement times, right hand (ms)	565	577	.821	-0.07
Total reaction and movement times, left hand (ms)	570	570	.658	0.13
Total reaction and movement times, grand mean (ms)	575	586	.987	0.00
Tracking Tests				
Sine tracking run 1 error (mm)	13.7	17.7	.120	0.46
Sine tracking run 2 error (mm)	8.1	9.8	.087	0.52
Random tracking run 1 error (mm)	7.3	9.9	**.019***	0.71
Random tracking run 2 error (mm)	7.5	9.1	.096	0.50
Arrows Perception Test				
Number of arrows correct	12.0	11.5	.064	-0.53
Divided Attention Test				
Tracking error (mm)	8.3	8.6	.336	0.28
Number of arrows correct	12.0	12.0	.847	-0.05
Omission of arrows response	0.0	0.0	1.000	-0.21
Complex Attention test				
Reaction time (ms)	414	450	.362	0.26
Movement time (ms)	293	276	.802	-0.07
Reaction time standard deviation (ms)	124	121	.380	0.25
Movement time standard deviation (ms)	44	42	.847	0.06
Number of lapse errors	0.0	0.0	.791	0.07
Number of invalid trials	0.0	0.0	.093	-0.61

Continued on following page

Continued from previous page

Test Measure	Median for Pass group (n=44)	Median for Fail group (n=16)	Mann-Whitney *U p*-value (***p<.05**)	Cohen-type effect size[1]
Planning Test				
Duration of positional faults (s)	5.9	5.3	.987	-0.01
Distance travelled (m)	4.9	4.7	.315	-0.29
Number of hazards hit	2.0	2.0	.823	-0.07
Number of crashes	1.0	1.0	.428	+0.23

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value.

6.2.3 Classification of On-Road Pass and Fail

Selection of variables that were offered to the BLR model followed the process described in Section 6.1.3.2. First, pragmatic decisions were made about the variables that would be excluded from being offered to the model. A list of these 29 measures along with the reasons for exclusion can be found in Appendix O.

The remaining variables were ranked by effect size. The next step was to examine the tolerance levels of the variables using the SPSS function ‘Collinearity diagnostics’. The effect sizes were ranked and the top 20 independent variables that were related to the dependent Pass/Fail outcome were checked for collinearity.

Three independent variables were deleted due to tolerance values <0.2 (see Section 6.1.3.2 for a description of tolerance testing and cut off values). The three deleted variables were age grouping, Random tracking run 2, and Complex Attention reaction time. Of the remaining 7 variables, the 12 variables with the highest effect size were chosen to be offered to model, with seven of those measures being from the *SMCTests* battery. In order of highest to lowest effect size these variables were: Trail Making Test B, Random tracking run 1, Trail Making Test A, Arrows perception number of arrows correct, Sine tracking run 2, Sine tracking run 1, Dementia Rating Scale -2 AEMSS, Age, Planning intersection safety margin, Visual Search number correct, Planning distance travelled, and Divided Attention tracking.

Table 6-6. Comparison of Pass and Fail on-road groups for normally distributed variables using *t*-tests.

Test Measure	Mean for Pass group (n=44)	Mean for Fail group (n=16)	*t*-test *p* value (***p*<.05*)**	Cohen's *d* effect size[1]
Years of driving	55.0	55.6	0.790	0.08
Driving Anger Scale time 1	33.2	31.8	0.609	-0.15
Big Five Inventory				
Extraversion	26.6	23.5	0.064	-0.55
Agreeableness	38.4	36.8	0.173	-0.40
Conscientiousness	37.3	34.8	0.136	-0.45
Neuroticism	18.5	18.1	0.740	-0.08
Openness to experience	35.1	35.6	0.793	0.08
Dementia Rating Scale -2 AEMSS[†]	11.1	10.1	0.205	-0.37
Ballistic Movement Test				
Peak velocity, right hand (ms)	951	972	0.717	0.11
Peak velocity, left hand (ms)	937	935	0.970	-0.01
Peak velocity, grand mean (ms)	944	954	0.857	0.05
Visual Search Test				
Mean reaction time (s)	4.8	5.0	0.390	0.26
Number correct	15.5	14.8	0.306	-0.30
Complex Attention test				
Total mean movement and reaction times (ms)	739	768	0.474	0.21
Planning Test				
Lateral road position error (mm)	2.7	2.7	0.569	0.17
Intersection safety margin (mm)	40.6	36.6	0.311	-0.30

[1]Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment.

The 12 measures were entered using a step-wise backwards elimination procedure. The model accepted one *SMCTests* measure – Random Tracking 1 – and one cognitive test measure – Trail Making Test B. These measures accounted for 25% of the variance in the on-road outcome (Nagelkerke R^2). The ROC AUC for the BLR model was .76 (z = 3.47; p <.001, 95% CI: .64–86). The sensitivities, specificities and total classification accuracies of the BLR across a range of cut-points are reported in Table 6-7.

Table 6-7. The sensitivities, specificities and overall classification accuracies of the BLR model at different cut-points including the 'Default', and 'Optimized' cut-points

Criterion[1]	Sensitivity (%)	Specificity (%)	Mean Sensitivity & Specificity (%)	Accuracy (%)
≥ 0.10	100.0	15.9	58.0	38.3
≥ 0.20	75.0	59.1	63.3	67.1
≥ 0.26 'Optimized'	68.8	75.0	71.9	73.3
≥ 0.30	62.5	79.6	71.0	75.0
≥ 0.40	43.8	88.6	66.2	76.7
≥ 0.50 'Default'	31.3	95.5	63.4	78.3
≥ 0.60	25.0	97.7	61.4	78.3
≥ 0.70	18.8	97.7	58.2	76.7
≥ 0.80	6.3	97.7	52.0	73.3
≥ 0.90	6.3	100.0	53.1	75.0

[1]Criteria displayed are the predicted probabilities of receiving a Fail score.

As described in Section 6.1.3.6, two cut-points were compared for each model. Using a default criterion value of ≥0.5 for detecting fails ('Default' cut-point), the model correctly classified 47 of 60 participants (78.3%) into on-road Pass or Fail groups, with a negative predictive value of 79.3%, and positive predictive value of 71.4%. The cut-point for the highest mean sensitivity and specificity value (mean = 71.9%, cut-point = 0.26, 'Optimized' cut-point) correctly classified 44 of 60 participants (73.3%) into on-road Pass or Fail groups with a negative predictive value of 86.8%, and positive predictive value of 50.0%.

The 60 iterations generated by leave-one-out cross-validation reduced the accuracy of the 'Default' cut-point from 78.3% to 70.0%, sensitivity from 31.3% to 12.5%, specificity from 95.5% to 90.9%, negative predictive value from 79.3% to 74.1%, and positive predictive values from 71.4% to 33.3%. Applying leave-one-out results to the 'Optimized' cut-point reduced overall accuracy from 73.3% to 60.0%, sensitivity from 68.8% to 43.8% (11 to 7 of the 16 Fails), specificity from 75.0% to 60.0%, negative predictive value from 86.8% to 76.3%, and positive predictive value from 50.0% to 31.8%. For a summary of these results see Table 6-8 .

Table 6-8. The sensitivities, specificities and overall classification accuracies of the BLR model at three different cut-points for classification models along with their accuracy following leave-one-out cross-validation

Model	Classification					Following leave-one-out cross-validation				
	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)
BLR										
‘Default’ cut-point	31.3	95.5	79.3	71.4	78.3	12.5	90.9	74.1	33.3	70.0
‘Optimized’ cut-point	68.8	75.0	86.8	50.0	73.3	43.8	65.9	76.3	31.8	60.0

Of the 60 iterations of the leave-one-out cross-validation, nine contained a different set of measures from Random Tracking run 1 and Trail Making Test B that were utilized in the overall classification model. Random Tracking run 1 appeared in all but one of the iterations, while Trail Making Test B was left out of 6 iterations. The Dementia Rating Scale -2 AEMSS was the most frequent measure to enter a model, with an appearance in 7 iterations. Divided Attention tracking was utilized in 3 iterations, Arrows perception number of arrows correct in 2, and Sine tracking run 1, Planning intersection safety margin, and Planning distance travelled in one iteration each. While most of the iterations employed two or three tests, one iteration included 8 measures and incorrectly predicted that a driver with an on-road Fail score would Pass.

Since only two variables entered the BLR model, the assessment could be completed in around 15 minutes and would require the use of one cognitive test (Trail Making Test parts A & B, with Trail Making Test part A being a necessary forerunner to part B) and two of the four Tracking Tasks (Sine Tracking run 1 and Random Tracking run 1, with Sine Tracking run 1 being a necessary forerunner and training test for the completion of Random Tracking run 1 which is included in the BLR model).

6.3 Discussion

6.3.1 Associations Between Independent Variables and On-Road Outcome

The independent variables with the largest effect sizes between Pass and Fail groups were the Trail Making Test B (Cohen-type effect of 0.80) and Random Tracking run 1 (0.71). The majority of effect sizes were in the small to moderate range (0.20 – 0.50) indicating that there were few measures that separated groups to a potentially useful degree. This is not particularly surprising given that the participants were screened to be ostensibly cognitively healthy.

6.3.2 Classification of On-Road Driving Ability

Using the ‘Default’ cut-point of 0.5, BLR utilized TMT B completion time and Random Tracking 1 to correctly classify 78.3% of the participants into on-road Pass or Fail groups. This value is only marginally higher than the rate that would have been achieved by predicting that every driver would Pass (44 passed, 73.3% of the sample). The ‘Optimized’ cut-point had a higher sensitivity but a lower specificity with a subsequently lower accuracy of 73.3%.

Random Tracking run 1, one of the tests selected by the BLR model, measures visuomotor planning and execution, with lower accuracy scores associated with an increased likelihood of an on-road Fail outcome. Random Tracking run 1 is performed after Sine Tracking run 1, which many participants find difficult initially. This is usually resolved by the end of the trial. Thus, Random Tracking run 1's ability to classify driving ability may reflect either difficulties with visuomotor control or with delayed learning of the tracking task that extends past the first tracking trial. The other test selected by the BLR model, Trail Making Test B, consists of visual scanning, sequencing, and task-switching, with greater time to completion associated with a Fail score. As Trail Making Test B is a sensitive detector of cognitive impairment, lower scores on this test could indicate the presence of undetected cognitive impairment.

As expected, the accuracy of the BLR model was reduced following leave-one-out cross-validation for both cut-points. The model suffered primarily in sensitivity for predicting Fails with specificity less affected. The drops in accuracy following leave-one-out cross-validation emphasises the importance of investigating models beyond classification in order to estimate their stability and likely performance in an independent sample.

To be used in a practical setting, considerations of the appropriate recommended cut-point would depend on factors such as the cost of more comprehensive driving assessment, and the percentage of Passes that would initially be flagged for further, unnecessary, testing. This information is provided by the negative and positive predictive values. The negative predictive value indicates the percentage of predicted Passes that would actually Pass the on-road assessment. Following leave-one-out cross-validation these values ranged between 74.1% and 76.3% for the two cut-points. This suggests that around 25.0% of examinees predicted to Pass would actually Fail the on-road assessment. The positive predictive values show the percentages of examinees predicted to Fail who would actually Fail the on-road assessment. The range of 31.8% to 33.3% depending on the cut-point indicates that around 67% of participants predicted to Fail would actually Pass the on-road assessment. It is clear that potentially testing so many predicted Fails with an on-road assessment only to find that two thirds are actually Passes would be an inefficient use of driving assessment resources for older drivers who have no diagnosed cognitive impairment and limits the usability of the model.

6.3.3 Generalization of the Results

There are several reasons why those participating in the study may not have come from the general population of older drivers. Firstly, many of the participants came from church groups, social groups, and exercise groups. Thus, drivers not participating in such social activities were less likely to be approached for participation. Secondly, many participants were recruited via word-of-mouth from participants already in the study. Once again, those with fewer social contacts will have been less likely to be asked to participate. Thirdly, on-road driving assessments can be anxiety provoking, and many drivers who had the research proposed to them may have declined to participate due to concerns about completing a driving assessment. Also, the study design was quite demanding, with participants required to attend an approximately three hour testing session and a one-hour driving assessment on separate days. It may be that those who participated in the study were more conscientious on average than the general population of healthy older drivers.

It is also likely that some of the participants were not cognitively healthy, especially since dementia prevalence has reported rates between 13% and 43% in the 80 to 89 age group, increasing exponentially per year within this age range (Ritchie & Kildea, 1995; De Ronchi et al., 2005; Plassman et al., 2007). Inclusion criteria for the study stated that participants did not have a diagnosed cognitive impairment. Some participants are likely to have had undiagnosed impairments due to problems not being discussed with a person's general practitioner, or due to sub-clinical deterioration. The researcher suggested that one participant mention the detection of possible memory impairments to his GP. He was subsequently sent for a cognitive assessment and diagnosed with dementia. This did not violate our inclusion criteria as he had no diagnosis at the time of the testing and, hence, would have been considered to be within the population of low-risk older drivers by his GP. The sample were subsequently followed for 24 months as detailed in Chapter 7 as part of the Healthy Driver Follow-up during which time three participants reported having a stroke, four reported having a head injury and one reported being diagnosed with Parkinson's disease (see Section 7.2.4 for details of reported medical conditions). It could be expected that Participants with subsequent strokes and Parkinson's disease and dementia may have been impaired at the time of testing for the current study (of these participants the three who had a stroke received a Pass score and the one with Parkinson's and one with dementia both received a Fail score).

However, once again the lack of diagnosis at the time of testing allowed their inclusion in the study.

Because of potential limitations in the recruiting process, the predictive value assigned to the individual tests used in the BLR models could be lower or higher for the general population of healthy older drivers.

6.3.4 Errors in Prediction of On-Road Assessment Outcome

There are several areas where errors in on-road prediction may have arisen. Firstly, it is likely that there are variables that have an impact on driving that were not measured in the study. These could include propensity to become anxious during the on-road assessment, the impact of personality traits such as sensation seeking and subsequent effects of risk-taking, and lack of road rules knowledge that could lead to the commission of errors that could affect the on-road driving outcome.

Secondly, there are differences in the ways that known or unknown medical conditions may have affected people's driving ability. Although the sample excluded people with known neurological conditions or moderate to serious head injuries, illnesses such as cancer and heart disease were common and, as discussed in Section 6.3.3, cognitive disorders may have been present yet undetected at the time of testing. The individual extent to which a medical illness may have affected a participant's driving performance cannot be known. Also, the abilities of participants to learn and perform cognitive and *SMCTests* are likely to be different. The skills involved in driving are well practised and to a large degree performed with little conscious effort, yet the fluid thinking required to learn and complete new tasks may vary between people. Some may have found it difficult to learn new tasks but have had no problem with the automatic skills associated with driving. Thus, difficulties with novel off-road tests may not correlate well with performance in the over-learned and mostly automatic task of driving.

Thirdly, the on-road driving assessment may itself have contained errors and, as detailed in Section 3.8, like most on-road driving assessment the Pass and Fail outcome was not based on pre-defined error lists and cut-off scores to decide whether a safe level of driving had been reached. There were also differences in the on-road driving conditions and experiences between participants. Although each person was taken on the same driving route, daily driving conditions varied from fine and sunny, to cold and raining. Assessments also occurred at different times of the day, and during different times of the year (August through December). The different road conditions encountered likely made some assessment drives more challenging than others.

Fourthly, BLR may not have been the best model for determining on-road Pass and Fail outcome. Perhaps a more computationally advanced model could have better taken into account subtle changes and interactions between variables that could result in a more accurate prediction. This possibility was tested in a recent study from our research group (Innes et al., in press) that investigated the accuracy of five non-parametric modelling techniques (BLR, nonlinear causal resource analysis, product kernel density, kernel product density, and support vector machine) along with one parametric measure (discriminant analysis) in utilizing *SMCTests* measures to predict on-road assessment Pass/Fail outcome in a group of 501 drivers referred due to brain disorders (163 suspected or probable dementia, 153 stroke, 113 traumatic brain injury, 27 Parkinson's disease, 9 brain tumours, and 36 other neurological disorder). The six classification models had ROC AUC values between 0.80 and 0.99, with product kernel density and support vector machine models correctly classifying over 99% of the group into Pass and Fail outcomes (41% of the sample failed the on-road assessment). Following leave-one-out cross-validation, the percentage accuracies of the product kernel density and support vector machine models dropped to the mid 70s, with the estimated predictive accuracies of all models falling between 71.0% and 75.8%. Thus, Innes et al. (in press) found that the classification accuracy of even advanced computationally complicated models performed similarly with more basic models such as BLR and discriminant analysis following leave-one-out cross-validation. It is therefore likely that the BLR model used in the current was similar in accuracy compared to other available modelling techniques.

Finally, the lack of consequences for poor performance could have influenced some drivers to drive with less care than they would have had the serious repercussions of a Fail outcome been possible. This could have caused a disconnect between generally good off-road testing scores and a poor on-road performance which belies the participant's actual ability to drive safely.

6.3.5 Comparison with Classification Models of Driving in the Literature

Few studies have investigated the driving ability of predominantly healthy older drivers in enough detail to derive levels of sensitivity and specificity for predicting a Fail or 'unsafe' driving score. Table 6-9 displays results from the current study as well as four independent studies, one of which investigated three different assessment batteries (Monash University Accident Research Centre, 2004). The studies contain a mixture of cognitive, physical and sensory-motor measures, with the Useful Field of View test utilized in the De Raedt and Ponjaert-Kristoffersen (2000), Stav et al. (2008) and Monash University Accident Research Centre (2004) studies. The Wood et al. (2008) study is the closest in design to the current study, with similar reported rates of sensitivity and specificity. The variance accounted for by the BLR model of .26 is also very close to the value for the current study of .25.

The Monash studies recruited New Zealand adults aged 80 and over who were in the process of completing an on-road driving assessment in order to renew their drivers licence. The sample would have included both cognitively healthy and unhealthy participants, and participants were not excluded due to medical conditions. The authors provide tables depicting the range of cut-off scores for each model with their respective overall accuracies, sensitivities, specificities, positive and negative predictive power and false positive and negative rates. The authors of this study highlighted the cut-points with the maximum balance of sensitivity and specificity, which are reported in Table 6-9. On average the Monash studies found much lower rates of overall accuracy as well as sensitivity and specificity compared to the Wood et al. and Healthy Older Drivers studies. Despite this, the authors of the Monash studies conclude that all three of the assessment batteries had strong relationships to the on-road performance of their participants and could be useful as screening tests within a licensing renewal context. No estimation of predictive accuracy was performed, so there is no way to tell how stable these models may be. The DriveABLE assessment battery was not able to be broken into its separate test measures and, hence, only the total score was offered to the regression model.

The De Raedt and Ponjaert-Kristoffersen (2000) and Stav et al. (2008) studies did not classify participants as 'safe' or 'unsafe' following their on-road assessment, but rather predicted the outcome of their rating on a continuous driving scale. Thus, sensitivities and specificities cannot be reported but both show a much higher values for variance accounted for (0.64 and 0.44 respectively) than the other studies in Table 6-9. In some part, this higher variance accounted for will be due to the variables being better able to utilize variance to model a continuous outcome rather than a dichotomous outcome. In both the De Raedt and Ponjaert-Kristoffersen and Stav et al. studies, the participant group comprised some people who had been referred for a driving assessment, suggesting that their sample likely contains participants at higher risk of unsafe driving compared to the Healthy Older Drivers and Wood et al. (2008) studies. Also, some participants in the Stav et al. study had MMSE scores below 24 indicating that some were likely cognitively-impaired In the Stav et al. study on-road driving assessors were not blinded to the off-road testing results which introduces bias into the on-road rating system.

Only the Wood et al. (2008) study went beyond classification in order to approximate how their model would generalize to a new sample as shown in Table 6-10 Leave-one-out cross-validation and testing the model on a holdout sample of 20% of participants both produced sensitivities and specificities very close to, and in some cases higher than, the levels found in their classification model. This is in contrast to the drops in sensitivity and specificity found in the Healthy Older Drivers classification models. This is likely due to the larger sample size of the Wood et al. study which led to the construction of a more stable classification model. It could also be that the measures that entered the Wood et al. model were truly more useful in prediction of driving ability than those used in the current study. Finally, the on-road assessment used by the authors could have been more reliable than the assessment used in the current study which could have increased the ability of predictor variables to form a more accurate model.

.

Table 6-9. Comparison of the current study's classification models to other classification models with predominantly cognitively healthy older adult drivers.

Reference	Statistical Model	Measures in Model	Accuracy	Sensitivity	Specificity	Variance Accounted for (R^2)	n	Mean Age of Sample
Healthy Older Drivers (the current study)	Binary logistic regression	Sensory-motor tracking task, Trail Making Test B	73.3%[1]	68.8%[1]	75.0%[1]	0.25	60	76.7
Wood et al. (2008)	Binary logistic regression	Reaction time, motion sensitivity, postural sway, self-reported kms driven per week	73.7%	91.0%	70.0%	0.26	270	75.8
Monash University Accident Research Centre (2004)	Binary logistic regression	Five measures from the GRIMPS[2]: rapid-pace walk, foot-tap test, delayed word recall, Trail Making Test B, visual acuity	62.1%[3]	70.1%[3]	58.4%[3]	N/S[4]	284	82.4[5]
Monash University Accident Research Centre (2004)	Binary logistic regression	Three measures from the CALTEST[6]: autotrails, Useful Field of View – divided attention, Useful Field of View – selective attention	67.6%[3]	56.3%[3]	71.5%[3]	N/S[4]	284	82.4[5]
Monash University Accident Research Centre (2004)	Binary logistic regression	Total score from the DriveABLE computerized test battery	59.9%[3]	65.3%[3]	57.9%[3]	N/S[4]	300	82.4[5]
De Raedt and Ponjaert-Krisoffersen (2000)	Multiple regression	Movement perception, Useful Field of View, cognitive flexibility, selective attention	N/S[4]	N/S[4]	N/S[4]	0.64	84	78.6
Stav et al. (2008)	Multiple regression	Contrast sensitivity, rapid-pace walk, Useful Field of View, MMSE total score	N/S[4]	N/S[4]	N/S[4]	0.44	123	75.3

[1]Percentages are given for the cut-point with the highest average of sensitivity and specificity. [2]GRIMPS = Gross Impairments Screening Battery of General Physical and Mental Abilities [3]Percentages are given for the cut-point with the highest average of sensitivity and specificity [4]N/S = not stated and unable to be computed from provided data. [5]The sample was a subset of a larger sample of 852 participants with a mean age of 82.4. [6]CALTEST = Department of Motor Vehicles, California test

Table 6-10. Comparison of the current study's estimated predictive accuracy compared to the estimation of predictive accuracy of other studies using samples of predominantly cognitively healthy older adult drivers

Reference	Method of estimating predictive accuracy	Measures in Model	Accuracy	Sensitivity	Specificity	n	Mean Age of Sample
Healthy Older Drivers (the current study)	Leave-one-out cross-validation of binary logistic regression model	12 measures were offered, with most iterations utilizing a sensory-motor tracking task and Trail Making Test B	60.0%[1]	43.8%[1]	65.9%[1]	60	76.7
Wood et al. (2008)	Leave-one-out cross-validation of binary logistic regression model	Reaction time, motion sensitivity, postural sway, self-reported km driven per week	73.7%	87.0%	71.0%	270	75.8
Wood et al. (2008)	Testing model against 20% holdout sample	Reaction time, motion sensitivity, postural sway, self-reported km driven per week	74.4%	92.0%	71.0%	270	75.8

[1]Percentages are given for the cut-point in the classification model with the highest average of sensitivity and specificity.

6.4 Review of Study Hypotheses

Given the results of the study, it is possible to address the hypotheses outlined in Section 4.1.

1. A combination of standard cognitive tests and *SMCTests* measures will provide on-road Pass and Fail accuracy statistics higher than that achieved by previous studies.

This hypothesis is not supported as the current study's classification model was lower than the accuracy of Wood et al. (2008), particularly following leave-one-out cross-validation. The BLR model in the current study performed better at classifying on-road assessment outcome than the Monash University Accident Research Centre (2004) studies.

The current study is not as easy to compare to studies of De Raedt and Ponjaert-Kristoffersen (2000) and Stav et al. (2008) since participant samples likely contained more impaired people in these studies. Because the authors of these studies used multiple regression rather than binary logistic regression, sensitivities and specificities could not be calculated, and measures of variance accounted for may be higher due to predicting a continuous rather than a dichotomous outcome.

2. Participants with a higher score on the Driving Anger Scale will be more likely to Fail an on-road driving assessment.

This hypothesis was not supported as neither administration of the Driving Anger Scale was significantly related to Pass and Fail groups, with effect sizes of 0.15 and 0.40 respectively. In any case, the direction of the effects were opposite from expected with the Pass group having higher scores than the Fail group.

6.5 Summary

Sixty drivers with no diagnosed cognitive disorder aged 70-84 years (mean age 76.7, 50% male), performed standard cognitive tests, computerized sensory-motor and cognitive tests (*SMCTests™)*, and measures of personality to form classification models of on-road assessment Pass and Fail outcome. Sixteen participants failed the on-road assessment. A backwards stepwise binary logistic regression model selected a measure of executive function and a computerized measure of visuomotor planning and coordination. Following leave-one-out cross-validation, this model was estimated to correctly predict 60% of an independent group of cognitively-unimpaired older drivers into on-road Pass and Fail groups.

The sensitivity of the model for detecting Fails at the Optimized cut-point following leave-one-out cross-validation was 43.8%, meaning that over half of those who Failed the assessment were not being detected. Also, the positive predictive value of 31.8% at this same cut-point shows that around 70% of people predicted to Fail using the model would actually Pass the on-road. These accuracy statistics are not high enough to recommend the use of the model produced for the Healthy Older Drivers sample to be used in primary health care as a screen for possible driving problems.

CHAPTER 7 -
Study 2 – Healthy Driver Follow-Up

The objectives of the Healthy Driver Follow-up study were to follow participants from the Healthy Older Drivers study for 24 months with annual interviews to collect information about driving behaviour, health, and the commission of crashes and traffic offences. We were interested in whether a number of measures had a relationship with future crashes or traffic offences. These measures were Pass or Fail outcomes on the on-road assessment, cognitive or sensory-motor measures, reported lapses of attention and errors taken from the Driver Behaviour Questionnaire, and distress associated with medical conditions.

7.1 Methods

Ethical approval was obtained for this study through the Upper South A Canterbury Ethics Committee.

7.1.1 Participants

The sample for the Healthy Driver Follow-up study was the same as that recruited for the initial Healthy Older Drivers study (see Section 6.1.1). In summary, at initial recruitment the sample was a convenience group of 60 participants (50% male) aged 70 to 84 with 10 men and 10 women in each of three age groups (70-74, 75-79, and 80+ years). Exclusion criteria included a self-reported history of moderate to severe brain injury, a diagnosed neurological or cognitive disorder, severe musculoskeletal disease, and acute psychiatric disorder. No participant had an SMMSE score of below 27, suggesting none had cognitive impairment at a level indicative of dementia at the study baseline. It is unknown how many of the participants were told of their on-road Pass or Fail status by the occupational therapist following the on-road driving assessment that was part of the Healthy Older Drivers study.

Participants were invited approximately 5 months following the final participant's on-road assessment to participate in the Healthy Driver Follow-up study which included annual ~30-minute telephone interviews. One participant refused to take part in the Healthy Driver

Follow-up study and one refused to give access to official crash and traffic offence records. This left 12-month interview data for 59 participants and crash or offence data for 58. At 24-months, one participant had died and one had moved and could not be located, leaving interview data for 57 participants and crash or offence data for 56. Of these 56 participants, 41 (73.2%) received a Pass and 15 (26.8%) received a Fail on the on-road assessment. Of the four participants without full data available, three had passed and one had failed the on-road assessment.

Permission was sought to access officially-recorded traffic offence data from the New Zealand Transport Agency and crash data from the Ministry of Transport. Participants gave informed consent, and received no compensation for their involvement in the study.

7.1.2 Assessment Procedure

The assessment procedure for participant involvement in the Healthy Older Drivers study is detailed in Section 6.1.2.

For the Healthy Driver Follow-up, participants were phoned at both 12 and 24 month anniversaries of their on-road assessment and both times completed two questionnaires (all tests are described in detail in Chapter 5). One questionnaire was a modified version of the Driving Habits Questionnaire (DHQ) which assesses a range of driving-related behaviours, and the second was the 24-item Driver Behaviour Questionnaire (DBQ) which assesses the 12-month frequency of 24 driving scenarios grouped into three subscales measuring lapses of attention, errors, and deliberate violations of road rules.

Officially-recorded crash data was provided by the Ministry of Transport recorded police-reported crashes. Traffic offence data provided by the New Zealand Transport Agency recorded offences issued in person by police officers (thus excluding fixed speed camera offences).

7.1.3 Data Analysis Methods

Because of the low base-rate of officially-reported crashes and the higher base-rates of self-reported crashes and officially- and self-reported traffic offences (see Section 2.5 for a detailed discussion) we determined that we would be unlikely to find significant differences between Pass and Fail groups for crashes alone. Thus, one or more instances of either a crash or an offence reported either by a participant or an official source over the entire 24-month follow-up period formed the crash/offence binary outcome variable.

Fisher's Exact Test, a test of significance used for contingency tables when the expected cells frequencies are low (Glantz, 2005), was used to investigate whether those who failed the on-road assessment were more likely to go on to have crashes or offences for both retrospective (official data only) and prospective (both official and self-reported) data. An odds ratio and confidence interval were also computed. The phi coefficient, a measure of association between two binary variables, was calculated to show the strength of the relationship between Pass and Fail outcome and the binary incidence of a crash/offence in the 24-month follow-up period. The phi coefficient can be interpreted in a similar way to that of a bivariate Pearson correlation coefficient, with a value of zero indicating no association between variables and with larger positive or negative associations indicating a stronger relationship. Cohen (1988) suggests some tentative guidelines for interpretation of the phi coefficient, with a value of 0.10 considered a small effect, a value of 0.25 considered a medium effect, and a value of 0.39 and above considered a large effect. For the purposes of the current study, an effect size would need to be large in order for the Pass and Fail results of an on-road test to be considered useful for determining the risk of future crashes and traffic offences. Even a large effect size of 0.39 includes much overlap between the two compared distributions, and if measures such as revocation of a drivers licence are to take place based on the outcome of an on-road assessment, it is preferable that these distributions overlap as little as possible.

DHQ responses were investigated for changes in driving practices between 12- and 24-month interviews using McNemar tests which are designed for non-parametric nominal repeated-measures designs.

Participants were grouped into crash/offence and no-crash/offence groups to investigate whether scores on baseline testing were related to subsequent crash or offence involvement. Shapiro-Wilk tests were used to determine if independent variables were normally-distributed, with *t*-tests used to compare normally distributed data and Mann-Whitney *U* tests for non-normally distributed data.

Mann-Whitney *U* tests were used to investigate whether the number of medical conditions, medications taken, or the amount that participants were bothered by their medical conditions were related to crash or offence involvement. The amounts that participants reported being 'bothered' by each medical condition were assigned an ordinal rating of 0 for 'not at all', 1 for 'a little', 2 for 'a great deal' and were summed.

Results of the DBQ were investigated using Mann-Whitney *U* tests to find associations between crash/offence status and self-reported errors, violations, and lapses. All tests were two-tailed and α was set at .05.

7.2 Results

7.2.1 Retrospective and Prospective Crashes and Traffic Offences by Pass and Fail Groups

There were no officially-recorded crashes for the five-year period prior to the on-road assessment but six participants had a speeding traffic offence recorded in this period. All six of these participants went on to receive a Pass score on the on-road assessment. There was no difference in the incidence of five-year retrospective offences between Pass and Fail groups (p = .32).

There were no officially-reported crashes for the 24-month follow-up period. Over 24 months, six participants self-reported crashes (4 of these participants Passed the on-road assessment and 2 Failed) and eleven participants self-reported at least one traffic offence (7 of these participants Passed the on-road assessment and 4 Failed). All officially-reported offences were also self-reported and all were for exceeding the speed limit. In total, 16 participants had self-reported crashes or self- or officially-reported offences over the 24-month period (10 participants with an offence only, 5 participants with a crash only, one participant with both). Of these, 11 had received a Pass score on the on-road assessment (26.8% of all participants who received a Pass score) and five had received a Fail score on the on-road (33.3% of all participants who received a Fail score) (see Table 7-1).

Table 7-1. Number of participants who passed and failed the on-road assessment and those who went on to have a self-reported crash or self- or officially-reported traffic offence over the following 24 months

	24-month crash or traffic offence			
On-road outcome	No	Yes	Totals	% who had a crash or offence
Fail	10	5	15	33.3%
Pass	30	11	41	26.8%
Totals	40	16	56	28.6%

The odds ratio for prospective crashes and offences was 1.36 (95% CI: 0.38–4.89), indicating that drivers who received a Fail on the on-road assessment were 36% more likely to experience a crash or traffic offence in the following two years, although this odds ratio was not statistically significant (p = .741). The phi coefficient (r_φ = 0.06) indicates little, if any, association between Pass and Fail status and subsequent crashes or offences.

7.2.2 Driving Habits Questionnaire and Additional Driving Questions

Responses on the DHQ and additional questions are summarized in Table 7-2. At the 12-month interview, one participant had not driven for five months due to an injury but had resumed by the 24-month interview. From the 12-month to the 24-month interview there was a reduction in the number of people self-rating their driving quality as 'average' (McNemar's test, p = .035) and a concomitant trend for more people to rate themselves as 'good' drivers. There was also a reduction in the number of participants reporting driving in places outside the South Island (p = .039) and in those reporting regular use of bicycles (p = .016) and walking (p = .039).

7.2.3 Baseline Test Scores and Subsequent Incidence of Crashes and Offences

The results of baseline off-road testing were compared between the 16 participants who had a crash/offence versus the 40 with no crash/offence (Table 7-3, Table 7-4, and Table 7-5). The only significant differences were that participants in the crash/offence group drove a higher average number of km at baseline (z = -2.20, p = .028) and had *lower* error scores on the *SMCTests* test Divided Attention Tracking (z = -2.54, p = .011).

7.2.4 Medical Conditions and Subsequent Crashes and Offences

Details of reported medical illnesses at 12 and 24 months are presented in Table 7-6. The top five medical conditions at both 12 and 24 months were arthritis, high blood pressure, high cholesterol, cataracts, and heart disease. Between those with and without crashes or offences there was no difference in the number of medical conditions endorsed at 12 or 24 months (z = -3.14, p = .753; z = -0.09, p = .927 respectively), no difference in the number of medications taken for illness at 12 or 24 months (z = -0.93, p = .352; z = -0.67, p = .502 respectively), and no difference in the amount that participants were bothered by their condition/s at the 12-month interview (z = -1.54, p = .124). However, those who reported they were bothered by a medical condition at the 24-month interview were more like to have a crash/offence (z = -2.01, p = .044), with a moderate Cohen-type effect size of 0.59.

Table 7-2. Number of people endorsing items on the driving habits questionnaire and additional driving questions at 12- and 24-month interviews

Questionnaire item	12-month follow-up (n=59)	24-month follow-up (n=57)	McNemar Test *p*-value (two-tailed, ***p*<.05**)[1]
Currently driving	58[2]	57	1.000
Wear glasses when driving	35	31	.146
Wear seatbelt when driving	59	57	1.000
Preferred way to get around:			
drive oneself	54	47	.146
have someone else drive	4	8	.227
use public transportation or a taxi	1	2	1.000
Speed of driving relative to other cars on the road:			
much faster	0	0	-
somewhat faster	5	3	.500
about the same	51	51	.687
somewhat slower	3	3	1.000
much slower	0	0	-
Has been told to limit or stop driving	0	0	-
Self-rated quality of driving:			
excellent	9	9	1.000
good	31	38	.134
average	19	10	**.035***
fair	0	0	-
poor	0	0	-
Likely action when a person doesn't feel like driving:			
ask a friend of relative to drive	33	33	.824
call a taxi or take the bus	18	17	1.000
drive oneself regardless of the situation	2	4	.687
cancel or postpone plans	6	3	.453
Types of driving performed in the past 3 months:			
driving when raining	58[3]	55	.500
driving alone	57[3]	57	1.000
parallel parking	53[3]	48	.219
making right-hand turns across oncoming traffic	56[3]	57	.500

Continued on following page

Continued from previous page

Questionnaire item	12-month follow-up (n=59)	24-month follow-up (n=57)	McNemar Test *p*-value (two-tailed, ***p*<.05**)[1]
driving on motorways or highways	54[3]	52	.625
driving on high-traffic roads	57[3]	57	1.000
driving in rush-hour traffic	55[3]	53	.625
driving at night	54[3]	53	1.000
People reporting accidents over 12 months	4	2	.687
People reporting accidents attended by police over 12 months	0	1	1.000
People reporting being pulled over by police over 12 months	4	2	1.000
People reporting receiving traffic tickets over 12 months	10	4	.180
People who during the last year have driven:			
in their immediate neighbourhood	59	57	1.000
beyond their immediate neighbourhood	58	57	1.000
to neighbouring towns	53	51	1.000
to distant towns	33	29	.791
to places outside of the South Island	10	2	**.039***
Forms of transport used regularly:			
drive own car	58	57	1.000
taken as passenger	57	53	.375
taxi	9	4	.180
motor scooter / motorcycle	1	0	1.000
bicycle	12	4	**.016***
bus	32	37	.146
walking	52	44	**.039***
Driving lessons in the last 12 months	0	0	-
Asked by their doctor about their driving in the last 12 months	12	6	.146

[1]The McNemar test is a repeated-measures test and calculations could only be performed for the 57 participants for whom both 12- and 24-month data were available. [2]One driver had ceased driving for several months at the 12-month interview due to a leg injury but had begun driving again by the 24-month interview. [3]One participant was excluded from tests of significance for these measures since they are repeated measures as, due to an injury, she had not driven during the 3 months before the 12-month interview.

Table 7-3. Comparison of non-normally-distributed baseline testing results for standard off-road tests between those with a 24-month crash/offence and those without

Test Measure	Median No-crash/offence group (n=40)	Median crash/offence group (n=16)	Mann-Whitney *U* *p*-value (two-tailed, ***p*<.05**)	Cohen-type effect size[1]
Gender (1 = male, 2 = female)	1.0	2.0	.315	0.30
Age (years)	78.0	81.5	.190	0.39
Handedness (1 = right, 2 = left)	1.0	1.0	.871	0.05
Years of education	13.0	15.0	.052	0.56
Occupation code (range 1–8, higher = more professional occupation)	2.0	2.0	.309	0.29
Km driven last 12 months	6747	11028	**.028***	0.65
Vision[2]				
Left eye	9.0	9.0	.206	0.38
Right eye	6.0	9.0	.346	0.28
Binocular	6.0	6.0	.992	0.00
Road sign test (no correct)	11.0	11.5	.930	0.02
Mini-Mental State Exam	29.0	29.0	.448	0.24
Geriatric Depression Scale	3.0	1.5	.276	0.31
Beck Anxiety Inventory	4.0	3.5	.956	0.02
Driving Anger Scale	31.0	31.5	.856	0.05
Trail Making Test A (s)	33.0	33.5	.696	0.13
Trail Making Test B (s)	89.0	93.5	.568	0.17
Wechsler Test of Adult Reading (estimated IQ score)	118.0	116.5	.080	0.49

Table 7-4. Comparison of non-normally-distributed baseline testing results for *SMCTests* between those with a 24-month crash/offence and those without

Test Measure	Median No-crash/offence group (n=40)	Median crash/offence group (n=16)	Mann-Whitney *U* *p*-value (two-tailed, ***p*<.05**)	Cohen-type effect size[1]
Footbrake and Clutch Test				
Mean reaction time (ms)	298.0	279.5	.072	0.52
Mean movement time (ms)	287.0	292.5	.906	0.03
Total reaction and movement times (ms)	581.0	572.0	.457	0.22
Ballistic Movement Test				
Reaction time, right hand (ms)	350.9	345.3	.913	0.03
Reaction time, left hand (ms)	337.2	334.3	.856	0.05

Continued on following page

Continued from previous page

Test Measure	Median No-crash/offence group (n=40)	Median crash/offence group (n=16)	Mann-Whitney *U* *p*-value (two-tailed, ***p*<.05**)	Cohen-type effect size[1]
Reaction time, grand mean (ms)	346.9	350.2	.828	0.07
Movement time, right hand (ms)	212.5	248.3	.147	0.42
Movement time, left hand (ms)	223.3	242.5	.420	0.23
Movement time, grand mean (ms)	221.6	242.4	.186	0.38
Total reaction and movement times, right hand (ms)	565.9	609.7	.446	0.22
Total reaction and movement times, left hand (ms)	569.7	596.8	.835	0.06
Total reaction and movement times, grand mean (ms)	570.2	606.0	.502	0.19
Tracking Tests				
Sine tracking run 1 error (mm)	15.8	13.0	.152	0.43
Sine tracking run 2 error (mm)	8.8	7.7	.063	0.59
Random tracking run 1 error (mm)	8.7	7.1	.446	0.22
Random tracking run 2 error (mm)	8.6	7.2	.301	0.32
Arrows Perception Test				
Number of arrows correct	12.0	12.0	.583	0.17
Divided Attention Test				
Tracking error (mm)	9.1	8.0	**.011***	0.82
Number of arrows correct	12.0	12.0	.576	0.16
Omission of arrows response	0.0	0.0	.527	0.22
Complex Attention test				
Reaction time (ms)	421.5	435.0	.568	0.18
Movement time (ms)	278.0	314.0	.280	0.32
Reaction time standard deviation (ms)	123.5	126.0	.683	0.13
Movement time standard deviation (ms)	41.0	40.5	.758	0.10
Number of omissions errors	0.00	0.00	.264	0.40
Number of commission errors	0.00	0.00	.430	0.22
Planning Test				
Duration of positional faults (s)	6.0	6.5	.856	0.05
Distance travelled (m)	4.9	4.7	.831	0.05
Number of hazards hit	2.0	2.0	.926	0.04
Number of crashes	0.5	1.0	.086	0.54

[1]Cohen's effect-size for rank-transformed variables (Hopkins, 2004). [2]Rated using a Snellen eye chart with a score of 9 equal to a metric 6/9 vision (20/30).

Table 7-5. Comparison of normally-distributed baseline testing results between those with or without a 24-month crash/offence

Test Measure	Mean no-crash/offence (n=40)	Mean crash/offence group (n=16)	*t*-test *p*-value (two-tailed, ***p*<.05**)	Cohen's *d* effect size
Years of driving	54.3	56.6	.304	0.31
Big Five Inventory				
Extraversion	25.5	26.5	.552	0.18
Agreeableness	38.9	37.7	.753	0.09
Conscientiousness	36.4	35.7	.692	0.12
Neuroticism	18.5	19.1	.660	0.13
Openness to experience	34.9	36.3	.422	0.24
Dementia Rating Scale-2 AEMSS[1]	10.7	10.9	.707	0.11
Ballistic Movement Test				
Peak velocity, right hand (ms)	980.9	895.6	.125	0.46
Peak velocity, left hand (ms)	949.0	902.2	.324	0.29
Peak velocity, grand mean (ms)	965.0	898.9	.187	0.40
Visual Search Test				
Mean reaction time (s)	4.8	4.7	.725	0.14
Number correct	15.0	16.3	.086	0.52
Complex Attention test				
Total mean movement and reaction times (ms)	740.5	779.7	.345	0.28
Planning Test				
Lateral road position error (mm)	2.7	2.6	.519	0.19
Intersection safety margin (mm)	38.7	40.6	.694	0.13

[1]AEMMS - age and education-adjusted MOANS scaled score (MOANS = Mayo Older American Normative Studies).

Table 7-6. People reporting medical illness, taking medication for illness, and reporting being bothered by illness at 12 and 24 month interviews

	12 months (n=59)			24 months (n=57)		
Medical Condition	People reporting illness	People taking medication	People bothered by their illness	People reporting illness	People taking medication	People bothered by their illness
Arthritis	32	12	12	32	10	13
High blood pressure	32	32	0	27	27	0
High cholesterol	22	18	0	22	18	0
Cataracts	18	0	2	22	0	5
Heart disease	16	14	2	18	18	1
Surgery	12	1	1	12	2	2
Cancer	12	3	3	8	2	0
Osteoporosis	9	8	1	10	8	1
Fall	6	2	4	7	1	3
Diabetes	6	5	0	6	5	0
Glaucoma	4	4	0	4	4	1
Thyroid problems	2	2	0	3	3	0
Anxiety	2	1	0	3	1	1
Macular degeneration	2	1	0	2	1	1
Stroke	2	2	0	1	1	1
Depression	1	0	0	6	4	2
Broken bones	1	0	0	4	1	2
Dementia	1	0	0	1	0	0
Sleep apnoea	1	1	1	1	1	0
Head injury	0	0	0	4	0	0
Parkinson's	0	0	0	1	1	1
Multiple sclerosis	0	0	0	0	0	0
Diabetic retinopathy	0	0	0	0	0	0
Retinal detachment	0	0	0	0	0	0

7.2.5 Driver Behaviour Questionnaire and Crash and Offence Involvement

There were no differences in the self-reported frequency of errors, lapses, and violations in drivers with crashes or offences, compared to those without, at either 12- or 24-months (see Table 7-7). The frequency of errors reported at 24 months was higher than the frequency reported at 12 months across crash/offence and no crash/offence groups combined (mean =

1.67 at 12 months, mean = 2.16 at 24 months, Wilcoxon signed-rank test, z = -1.99, p = .047) with no change in the frequencies of reported lapses or violations.

Table 7-7. Self-reported error, lapse and violation rates on the Driver Behaviour Questionnaire at 12 and 24 months by crash/offence group

Driver Behaviour Questionnaire composite measure	Median no crash/offence group (n=40)	Median crash/offence group (n=16)	Mann-Whitney *U* *p*-value (two-tailed)	Cohen-type effect size[1]
12-month errors	1.0	1.0	0.86	0.05
12-month lapses	4.0	4.0	0.92	0.03
12-month violations	1.0	2.0	0.70	0.12
24-month errors	2.0	2.0	0.81	0.07
24-month lapses	5.0	5.0	0.81	0.08
24-month violations	2.0	1.5	0.93	0.02

[1]Cohen's effect-size for rank-transformed variables (Hopkins, 2004).

7.3 Discussion

7.3.1 Association of Measures with Real-World Adverse Driving Events

Pass or Fail outcome on the on-road driving assessment was not related to either five-year retrospective officially-reported traffic offences or two-year prospective self- or officially-reported crashes and traffic offences. Effect sizes would have had to be of large magnitude to detect a significant difference in crashes or offences between the on-road Pass and Fail groups. However, large effects would also be required for the distributions of crash/offence and no crash/offence groups to be separated enough to support recommendations for licence cessation following the failure of an on-road assessment. This notwithstanding, the very small phi coefficient of 0.06 indicates that there was little if any association between whether a participant received a Pass or Fail on the on-road assessment and whether they went on to have a crash or offence in the following 24 months.

The New Zealand study by Keall and Frith (2004a) that was described in Section 3.8 showed a significant relationship between failing an on-road driving assessment and having an increased risk of future crashes but like the current study the phi coefficient was small. Compared with the Keall and Frith study, the Healthy Driver Follow-up study had more liberal criteria for defining a negative driving outcome as it included traffic offences as well as crashes, and allowed for self-reported events rather than officially reported events only.

Keall and Frith investigated only crashes recorded by the Ministry of Transport's crash database, which are crashes that culminate in injury or death. In fact, there were no officially-recorded crash events in the five-year retrospective or two-year prospective Healthy Driver Follow-up study. Unsurprisingly, the current study had a much higher incidence rate for the dependent variable (33% crashes of offences in the Fail group) compared to the Keall and Frith study (1.2% crashes in the Fail group). Even with the more sensitive criterion for defining adverse driving events the phi coefficient of 0.06 for the Healthy Driver Follow-up study was still very small.

The occurrence of a crash or offence over the following 24-month period was related to a higher number of km driven per year measured at the baseline testing session. Higher km driven per year is likely related to increased driving exposure which led to a greater likelihood of experiencing a negative on-road outcome (provided those who reported higher km at the initial testing session maintained higher km over the 24-month period). This tendency for drivers with higher reported driving distance to be more likely to have an adverse driving event has been found in some other studies (Owsley et al., 1998; Ball et al., 2006) but, conversely, many other studies have found that those older drivers driving fewer km were more likely to have an adverse driving event (Janke, 1991; Hakamies-Blomqvist et al., 2002; Keall & Frith, 2004b; Langford et al., 2006; Alvarez & Fierro, 2008). This finding of lower km in older drivers relating to a higher likelihood of adverse events is usually attributed to a core group of older drivers who are affected by cognitive or physical illness and limit their driving, yet still manage to have adverse driving events.

The second association that a *poorer* error score on the *SMCTests* Divided Attention Tracking task was related to increased crash or offence risk was unexpected and is more difficult to explain. It is possible that the result is simply a chance finding due to the number of comparisons performed (Type I error). The score pertained to *SMCTests* measure Divided Attention Tracking which requires participants perform two tasks simultaneously. The lower error score on the Random Tracking task component was not associated with a higher error score on the Arrows Perception task component which indicates that participants were not simply ignoring one task in order to perform well on the second. It is possible that those who scored well on the Divided Attention Tracking task had better physical or cognitive health which resulted in less cautious and more confident driving, and thus a higher likelihood of having a crash or offence. As no other baseline cognitive or sensory-motor tests were significantly associated with crash or offence status this hypothesis would seem unfounded.

The relationship between reports of ‘bother’ caused by medical conditions at 24 months and increased crashes or offences suggests that subjective physical or emotional discomfort associated with medical illness could affect a person’s likelihood of having an adverse driving event. There was an increasing trend in the number of broken bones, head injuries and depression between the 12- and 24-month interviews which could indicate declining physical and emotional health. The most common eye-related medical condition was cataracts with 22 people reporting it at 24-months, with glaucoma coming in a distant second with 4 people reporting it at both 12- and 24-month interviews. No relationships were found between crashes or offences and the number of medical conditions endorsed or medications taken. It is likely that knowledge of specific types of medical conditions, medications, or groupings of conditions, such as eye conditions, are more important than merely tallying the number of occurring medical conditions. The sample size in the Healthy Driver Follow-up study was too small to investigate these individually.

Changes in self-reported driving behaviour from the DHQ indicate that drivers reduced their exposure to long car trips over the 24-month follow-up period. This could be related to an overall reduction in driving confidence, although more participants rated themselves as ‘good’ drivers and fewer as ‘average’ at the 24-month compared to the 12-month interview. Decreased rates of walking and cycling could indicate declining physical condition which could also contribute to a reduction in long distance driving.

Finally, there was a significant increase in the frequency of reported errors on the DBQ from the 12-month to 24-month interview. Since the number of recalled violations and lapses did not also increase from 12 to 24 months, the increase in errors is unlikely to be due to increased vigilance following questioning at the 12-month interview in order to provide more accurate information at the 24-month interview. Therefore, a greater frequency of errors could be a real change with age. The fact that no relationships were found between the number of reported lapses, errors, or violations and the presence of crashes or offences indicates that the DBQ is not a promising screening tool for detection of older drivers at risk for adverse driving events.

7.3.2 Generalization of the Results

Since the sample for the Healthy Driver Follow-up study comprised the same participants as the Healthy Older Drivers study, the same limitations to generalization apply (see Section 6.3.3). Due to the limited amount of input required from participants in order to be part of the Healthy Driver Follow-up study, participation was high with only one participant declining participation. Similarly only one participant declined access to official records of crashes and traffic offences, and only two participants were not able to be contacted for the 24-month interview. This allowed for 56 of the original 60 participants to give full interview and official report outcome data. Although at the time of recruitment for the Healthy Older Drivers study all participants were without diagnosis of cognitive disorders, over the course of the 24-month follow-up two people experienced a stroke, one was diagnosed with dementia, and one with Parkinson's disease. Thus the sample in this study is perhaps more representative of a general older driver population than the sample had been at the commencement of the Healthy Older Drivers study.

Despite this, because of potential limitations in the initial recruitment process, the outcomes of the driving questionnaires and measures of adverse driving outcome could be lower or higher for the general population older drivers.

7.3.3 Practical Application of Results

Results of the five-year retrospective and 24-month prospective follow-up incidence of crashes or offences showed no significant association with a Pass or Fail score on an on-road assessment. This result, coupled with the null findings of the Anstey et al. (2009) study for associations between on-road Fail scores and 12-month self-reported crashes and the significant yet very small effect size of the Keall and Frith (2004a) study, suggests that on-road assessments are not useful for determining which of a group of cognitively healthy older drivers go on to have an adverse driving event in the following few years.

Results of baseline testing did not support hypothesized relationships between lower cognitive test scores and an increase in subsequent crashes and traffic offences. Thus, cognitive tests used for prediction of on-road adverse driving events in older drivers without a diagnosis of neurological disorder or cognitive impairment is not supported. It is, however, important to remember that an individual's self-report of cognitive impairment is not always accurate, and GPs are encouraged to screen their older patients for cognitive impairment when they present for licence renewal at ages 75, 80 and biennially thereafter. GPs are advised to use standardized and validated assessment measures to identify cognitive impairment coupled with pragmatic history-taking for driving ability.

The relationship between increased levels of bother caused by medical illnesses at 24-months and increased rates of crashes and offences may suggest that emotional or physical distress can impact on driving. For example, Molnar et al. (2007) found that being “bothered a great deal by diabetes mellitus” was related to motor vehicle crashes in older drivers. While in the Healthy Driver Follow-up study there was a much larger list of medical conditions investigated, the association with reports of bother and subsequent crashes and offences is intriguing and, provided this finding is found to be replicated in other research studies, could prove to be useful to general screening of possible problems with driving in the older population.

Reports of errors, lapses, and violations in the Driver Behaviour Questionnaire were not associated with subsequent crash/offence outcome, and although more errors were reported at the 24-month compared to the 12-month interview, strong evidence does not exist as to the usefulness of the measure in detecting older drivers who may go on to have an adverse driving event.

7.4 Review of Study Hypotheses

Given the results of the study, it is possible to address the hypotheses outlined in Section 4.2.

3. Drivers who fail an on-road driving assessment will have a greater incidence of crashes and traffic offences over the following years than those who pass.

This hypothesis is not supported as drivers who failed the on-road assessment did not have a significantly higher rate of crashes or offences in the following 24-months. Also, the effect size of r_{φ} = 0.06 was very small which indicates that the distributions of people with crashes or offences and no crashes or offences overlap enough to make separating them by Pass or Fail status impractical and inaccurate.

4. Poorer performance on standard cognitive measures are associated with increased crashes and traffic offences over the following years.

This hypothesis is not supported as no standard cognitive test results were significantly associated with increased incidence of crashes of offences over the 24-month follow-up.

5. Poorer performance on *SMCTests* measures are associated with increased crashes and traffic offences over the following years.

This hypothesis is not supported as only one *SMCTests* measure, Divided Attention Tracking, was associated with 24-month incidence of crashes or offences, and this association was in the opposite direction to that predicted, with a better score associated with an increased crash or offence rate. There are also several other instances in which scores were in the opposite direction to that expected, although these differences did not reach statistical significance, including smaller errors on Sine and Random Tracking tests, and faster reaction times and a higher number of correct trials on the Visual Search test. Other differences were small to non-existent.

6. Drivers who report a higher number of driving lapses or errors on the Driver Behaviour Questionnaire over the immediate years following an on-road driving assessment will be more likely to have had a crash or traffic offence.

This hypothesis is not supported as there were no differences in the number of lapses recorded and whether drivers had a crash or offence in the 24-month follow-up period. Parker et al. (2000) found an increase in lapses in a sample of drivers aged 49 and over compared to younger age groups. They also found that people with higher reported numbers of lapses were more likely to have had a crash in the previous five years, whether they caused the crash or it was due to the actions of another driver. Although the current study found no association with frequency of self-reported errors and lapses, the frequency of errors across the entire sample increased significantly from the 12- to the 24-month interview while there was no significant increase in the frequency of violations or lapses.

7. Drivers who report experiencing more distress associated with medical conditions are more likely to have crashes or traffic offences over the following years.

This hypothesis is supported as participants who reported being bothered by a medical condition or conditions at the 24-month follow-up interview (albeit not the 12-month interview) were more likely to have a crash/offence in the same period, with a moderate effect size of 0.59. This outcome could indicate that the presence of a bothersome medical condition can have adverse effects of a person's ability to drive safely.

7.5 Summary

Fifty-six participants from the Healthy Older Drivers study were followed for 24 months using annual telephone interviews to assess driving behaviour, driving attitudes, medical conditions, and the occurrence of crashes and receipt of traffic offences. Official data regarding crashes and traffics offences were also obtained. The aim was to determine whether on-road Pass/Fail classification could predict the incidence of crashes and traffic offences over two years and whether off-road measures performed during the Healthy Older Drivers study could predict subsequent crashes and offences. Failing the on-road assessment did not result in higher crash or offence rates and there were only two baseline measures that predicted crashes or offences (i.e., distance driven at baseline testing and a paradoxically lower error score on a measure of visuomotor planning and coordination). However, drivers who reported more distress associated with their medical condition(s) were more likely to have had a crash or offence at 24 months. The outcome of the Healthy Driver Follow-up study suggests there is little value in off-road or on-road assessment of older drivers without diagnosed cognitive impairment as these measures were not associated with future crashes or traffic offences.

CHAPTER 8 -
Study 3 – Dementia and Driving

The objective of the Dementia and Driving study was to find cognitive, sensory-motor, demographic, and medical factors that predict on-road driving performance in drivers with MCI or Alzheimer's dementia. A sample of 60 participants completed *SMCTests* measures, while a subset of 32 participants also completed a more extensive battery of cognitive tests. Binary logistic regression (BLR) was used to produce classification models, with leave-one-out cross-validation used to estimate the generalizability and stability of the models.

8.1 Methods

Ethical approval was obtained for this study through the Upper South A Canterbury Ethics Committee.

8.1.1 Participants

Participants were 60 referrals to the Driving and Vehicle Assessment Service (DAVAS) at Burwood Hospital, Christchurch who had diagnosed or suspected Alzheimer's dementia, mild cognitive impairment, unspecified cognitive impairment, or memory problems. Participants had a current full driver's licence, had not completed *SMCTests* testing at Burwood Hospital previously, and had no need for driving adaptations on their vehicle (such as a wheel spinner). Participants with a history of stroke, transient ischemic attack, moderate to severe head injury, multiple sclerosis, or any specific dementia other than Alzheimer's (e.g., frontotemporal dementia, dementia with Lewy bodies) were excluded from recruitment. The information sheet was mailed to participants and is provided in Appendix P.

There were two paths for participant involvement. The first was to complete approximately 5 hours of extra testing which included the completion of the extended test battery (see Section 5.7 for details). This testing was split across two testing sessions of around 2 to 2.5 hours length and was performed in the participant's own home on two separate days prior to the on-road assessment at Burwood Hospital. The second recruitment path consisted of completing the standard off-road (*SMCTests*) and on-road assessment required as part of the driving assessment referral and allowing the results to be included in the study. If a participant expressed interest in completing the extended testing, a family member or close friend was contacted and asked for information about the participant's completion of activities of daily living (see Section 8.1.2 below for details). A flowchart of recruitment is shown in Figure 8-1.

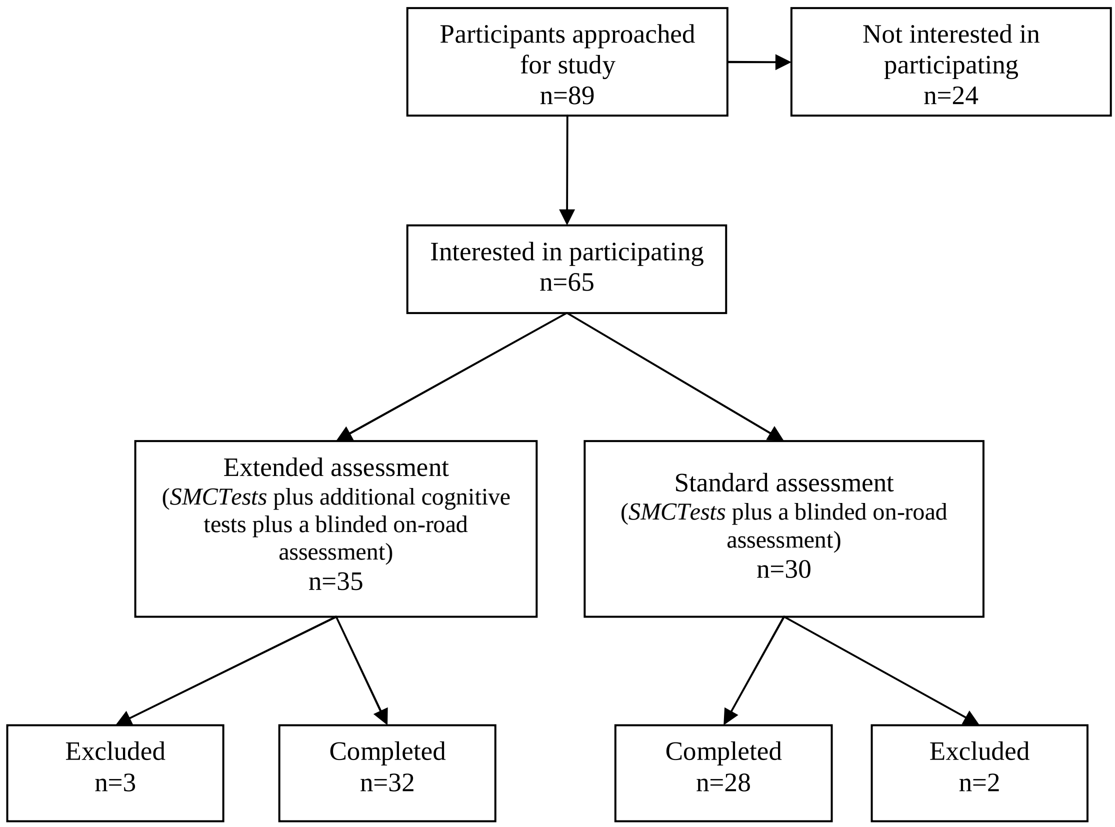


Figure 8-1. Numbers of participants approached, included, and excluded from the study.

Reasons for the twenty-four participants who declined to participate in testing included the intention to discontinue driving, missing off-road appointments and eventually having to be tested only with the on-road assessment and thus not completing *SMCTests* assessment, and being too upset about the necessity of the driving assessment to participate in the study. Thirty-five participants agreed to extended testing, with three subsequently excluded from the analysis. Of the three excluded, one was relocating and decided he would prefer to do his driving assessment in his new town, one was confirmed by a family member to have suffered two serious head injuries, and one was confirmed as not having dementia or MCI. Of the 30 people who agreed to have the results of only the standard assessment included in the study, 28 completed testing and two were excluded from the analysis. Of the two excluded participants, he spouse of one reported a history of transient ischemic attacks, and one participant was found to have frontotemporal dementia.

The final full sample consisted of sixty participants (36 males and 24 females) aged 58 to 92 (mean=77.9 years); 98% identified their ethnicity as New Zealand European. All had current driving licences, although some were not currently driving as they had recently been instructed to stop pending the results of the driving assessment. Fifty-six participants completed the assessment in a familiar vehicle, with the four remaining participants driving DAVAS's dual-control car. The subset of participants who completed additional cognitive testing consisted of 32 participants (22 male and 10 females) aged 58 to 92 (mean age = 76.2 years); 97% identified their ethnicity as New Zealand European. Thirty participants completed the assessment in a familiar vehicle, with the two remaining participants driving DAVAS's dual-control car. Their average years of driving experience was 55 years, with no difference in years of driving experience between males and females (male mean=56.5 years, female mean=50.7 years, z = -0.713, p = .476). Participants who completed the extended assessment (n=32) were compared to participants who completed only the standard assessment (n=28) on age and sex, which were the only demographic details (besides handedness) that were available for those who completed the standard assessment only. There was no difference in sex between groups with 69% of the extended assessment group being male as opposed to 50% in the standard assessment group (z = -1.467, p = .142), and there was an almost significant effect of age, with 76.2 years the mean age in the extended assessment group versus 79.9 years in the standard assessment group (z = -1.952, p = .051). All participants who failed the on-road assessment had a recommendation of licence revocation made to their referrer by the occupational therapist, as is standard practice for medical driving assessments.

8.1.2 Assessment Procedure

All participants completed an off-road (*SMCTests*) and on-road driving assessment appointment, with the *SMCTests* off-road assessment performed by the primary researcher in a session lasting around 70 minutes. On-road assessments were conducted an average of 6.0 days (SD = 4.3, range 0-28 days) after the off-road assessment. On-road assessments were administered by an experienced driving occupational therapist and a driving instructor who were blinded to off-road test performance. Details of the on-road assessment are described in

Section 5.10 including details of the on-road error checklist which was completed by the occupational therapist and driving instructor independently for each participant following the on-road assessment. The on-road assessors were not told of the medical conditions of participants beyond information available on the referral form which would have been variable between participants. Twelve participants started their on-road assessment from their own homes, with the remaining 48 participants beginning the assessment from Burwood Hospital. Participants completing their on-road assessment from home did so for a variety of reasons including difficulties in arranging for a family member to attend a Burwood Hospital appointment with them in case they failed the assessment and needed to be driven home, memory problems or confusion which made it too difficult to schedule and keep appointments, or assertion from participants or their families that they should be tested in their local area since that is the area where they perform their daily driving.

Extended assessment participants completed additional cognitive assessments in their own home over two appointments. Many participants had a spouse or another family member present at the beginning of the first appointment to aid in the collection of demographic and health information. Information regarding activities of daily living (see Section 5.9 for measures used) was completed by a significant other (family member or friend). For 22 participants, the informant was a spouse, for eight an adult child, and for two a close friend. All informants had known the participant for at least ten years and all lived in the same city and were in regular contact with the participant. Completion of the activities of daily living measures was conducted either at the end of the second testing session or by telephone.

For the subset of participants who completed extended cognitive testing, consent was sought to obtain officially-recorded data for traffic offences from the New Zealand Transport Agency and crashes from the Ministry of Transport for the five-year period prior to their on-road assessment. We recorded the number of demerit points earned in this time period as the independent variable for traffic offences.

8.1.3 Diagnosis of Alzheimer's Dementia and Mild Cognitive Impairment

A diagnosis of Alzheimer's dementia or mild cognitive impairment was determined for the 32 participants following extended cognitive testing. The NINCDS-ADRDA (McKhann et al., 1984) criteria were used for the diagnosis of Probable Alzheimer's dementia and

Petersen's (2004) criteria were used for the diagnosis of MCI. NINCDS-ADRDA criteria are reproduced below:

1. Dementia established by clinical examination and confirmed by neuropsychological tests.
2. Deficits in two or more areas of cognition.
3. Progressive worsening of memory and other cognitive functions.
4. No disturbance of consciousness (i.e., not present solely during the course of delirium).
5. Onset between ages 40 and 90, most often after age 65.
6. Absence of systemic disorders or brain diseases that, in and of themselves, could account for the progressive deficits in memory and cognition.

In addition to these criteria a significant deficit in activities of daily living was required for a diagnosis of Alzheimer's dementia (see below for details of how a deficit was determined).

Cognitive domains were defined as orientation, memory, language, praxis, attention, visual perception, and problem-solving (interpreted for this study as executive function) (McKhann et al., 1984). Cognitive domains were assessed with at least one specific cognitive measure (except language for which no cognitive measure was used) (see Table 8-1). For the measures for which standardized scores were available the cut-off for impairment was defined as a score less than or equal to the 5th percentile which is equivalent to a z-score of -1.64 and a Wechsler scaled score of $\leq$5.

A concerted effort was made to locate standardized data for ADAS-Cog subtest scores in order for percentile cut off scores to be determined to aid in MCI and Alzheimer's dementia diagnosis. Three studies were found that provided data for cognitively-unimpaired older adults (Zec et al., 1992; Graham et al., 2004; Grundman et al., 2004), but none provided means and standard deviations of subtest scores stratified by age. Assistance with normative data was requested from the Alzheimer's Disease Cooperative Study website (http://www.adcs.org/) and from the author (Dr. Richard Mohs) but was not forthcoming. Therefore for the two ADAS-Cog measures a pragmatic cut-off was chosen to reflect what was considered to be a cognitive deficit.

The criteria for MCI require a specific cognitive complaint, preferably corroborated by an informant, objective measurement of impairment (usually memory impairment), essentially preserved global cognitive functioning, largely intact activities of daily living, and no diagnosis of dementia. The same cut-points for scoring cognitive impairment for dementia were utilized for a diagnosis of MCI. The primary defining criterion was there was no significant deficit in functional activities. In an attempt to standardize this decision, a cut-off score of ≥ 3.44 from The Informant Questionnaire on Cognitive Decline in the Elderly was used, since a previous study found this point had a sensitivity of 100% for detecting dementia in clinical samples with a specificity of 86% (Jorm, 2004) (see Section 5.9.3 for more details on this questionnaire). Participants could be diagnosed as any of the four categories of MCI proposed by Petersen (2004): Amnestic MCI single domain, Amnestic MCI multiple domain, Non-Amnestic MCI single domain and Non-Amnestic MCI multiple domain (see Section 2.3.2 for more details on these diagnoses).

Table 8-1. NINCDS-ARDRA cognitive domains and the measures used to assess them for a diagnosis of Probable Alzheimer's dementia

Cognitive domain	Cognitive measure	Cut-off criteria
Orientation	ADAS-Cog Orientation subtest	≤4 correct responses (out of 8)
Memory	ADAS-Cog Delayed word recall subtest	≤4 recalled words (out of 10)
	Rey Complex Figure 3 min recall	≤5th percentile
Language	N/A	N/A
Praxis	Block Design	Scaled score ≤5
Attention	Letter-Number Sequencing	Scaled score ≤5
	Trail Making Test A	z-score ≤ -1.64
Visual perception	VOSP Incomplete Letters	z-score ≤ -1.64
	VOSP Silhouettes	z-score ≤ -1.64
Executive function	Colour-Word Interference, interference trial	Scaled score ≤5
	Letter Fluency	Scaled score ≤5
	Category Fluency	Scaled score ≤5
	Trail Making Test B	z-score ≤ -1.64

For those participants with a diagnosis of Alzheimer's dementia, an approximate severity score was based on the standardized version of the Mini-Mental State Exam. Deficit ranges as proposed by Mungas (1991) were used, with a score of ≤9 indicating severe impairment, 10-20 indicating moderate impairment, 21-24 indicating mild impairment, and 25 and above falling within the intact range. As discussed in Section 5.7.1, the MMSE is not a sensitive measure for detecting dementia, with many people with MCI and mild dementia likely to receive a score in the 'intact' range.

8.1.4 Data Analysis Methods

Data analysis methods were similar to those employed in the Healthy Older Drivers study (Chapter 6) with a few changes. Due to having a subset of participants complete extended cognitive testing, there were, in essence, two samples for analysis. The larger sample of 60 participants provided scant demographic information (sex, age, and handedness) and *SMCTests* results and were not diagnosed with either MCI or Alzheimer's dementia since they did not complete sufficient testing to perform a diagnosis. The subset of 32 participants who completed extended cognitive testing provided additional demographic information (i.e., years of driving, years of education), health information, and standard cognitive measures. For each of these samples, descriptive data were compared between on-road Pass and Fail groups and classification models formed using BLR, which was then tested for stability and generalizability to a new sample using leave-one-out cross-validation.

There was a small amount of missing data. In regards to *SMCTests*, two people were not administered the Planning test and one person did not complete the Divided Attention test. The tests were not administered due to the testing session running overtime. Two participants had incomplete data for the Complex Attention test due to the incorrect way in which they performed the task which provided no values for the reaction time and movement time measures. For the 32 participants who completed additional testing, one participant had no data for any of the Colour-Word Interference subtests due to red-green colour-blindness that prevented performance of the test. One person did not complete the Rey Complex Figure and one person did not complete the Finger Naming subset of the ADAS-Cog both due to oversights during test administration.

Missing data was not replaced by substitution of the mean but rather by substituting values through comparison on how the participant performed on a related test. This was done by finding the measure which best correlated with the measure for which the participant had missing data. The participant's score on the other measure they had completed was then compared to another participant with a similar score on the measure, and that new participant's score for the test with missing data was then substituted. In this way we hoped to provide a better estimate of the score the participant may have achieved if they had completed that measure.

8.1.4.1 Testing Normality of Test Data

The normality of the distribution of the data was assessed using the Shapiro-Wilk W test. Test data that was non-normally distributed was examined using non-parametric statistics, such as the Mann-Whitney *U* test, while normally distributed data was examined using parametric statistics such as *t* tests.

8.1.4.2 Binary Logistic Regression

The process of constructing the binary logistic regression (BLR) models was the same as for the Healthy Older Drivers study. The main concerns were to minimize the influence of overfitting on the model due to having too many variables to enter in relation to the sample size (a more detailed rationale for each of these steps is found in Section 6.1.3.2):

1. Variables pragmatically considered not to be useful in offering to the model were deleted from the variable list (i.e., variables that would not be accepted for use in a driving assessment service, variables with effects in the opposite from expected directions, and subtests bypassed in favour on more generalizable total scores).
2. Variables were ranked by effect size.
3. Variables with high collinearity with one or more other variables were deleted.
4. Remaining variables were offered to the model in order of effect size at the ratio of 1 variable per 5 participants (6 variables for n=32 extended assessment sample and 12 for n=60 standard assessment sample).
5. A model was formed using logistic regression with a backwards elimination procedure (criterion for entry to model p = .05, removal from model p = .10).
6. BLR models were tested for stability and generalizability using leave-one-out cross-validation.

Models were constructed in the manner described above for both the n=60 standard assessment sample and the n=32 extended assessment subsample.

We also examined whether including a balance of *SMCTests* measures, cognitive measures, and medical measures to enter the model would improve the fit and generalizability compared to the practice of offering the independent variables with the highest effect sizes to the model. It was determined that the two *SMCTests* measures with the highest effect sizes, the two cognitive measures with the highest effect sizes, and the two medical illnesses with the highest effect sizes would be offered to the BLR model. The two measures from cognitive and *SMCTests* had to be from different subtests in order that two separate measures from each domain would be entered rather than two measures from a single test. This may result in a more stable model than including two measures from the same test. Therefore, the n=32 extended assessment sample was used to construct two models, one according to the six steps outlined above (from here on referred to as “extended assessment model 1”), and one according to the amended criteria above (from here on referred to as “extended assessment model 2”).

8.1.4.3 Receiver Operating Characteristic (ROC) Curve

ROC curves were plotted for each classification model to determine if the classification was significantly better than random assignment and to compare the classification of the two n=32 models.

8.1.4.4 Leave-One-Out Cross-Validation

The process of leave-one-out cross-validation is detailed in Section 6.1.3.5 and was used to determine the stability and potential generalizability of the classification models.

8.1.4.5 Choice of Cut-Points for Reporting Accuracy

As with the Healthy Older Drivers study, two cut-points were examined for each model (see Section 6.1.3.6 for a detailed rationale). Cut-points used were (i) the default for Fail of 0.5, and (ii) the point with the highest average value of sensitivity and specificity found for each classification model. Again, we were also interested in the negative predictive value (false negatives) and positive predictive value (false positives) of each cut-point. The negative predictive value represents the proportion of all participants predicted to Pass who were actual Passes and the and the positive predictive value represents proportion of all participants predicted to Fail who were actual Fails. Avoiding unnecessary false positive and negative errors is important and will taken into consideration alongside sensitivity and specificity statistics.

8.1.4.6 Relationship Between the Driving Error Checklist and On-road Assessment Outcome

Intra-class correlations were computed using Cronbach’s α to assess reliability between the on-road error ratings provided by the occupational therapist and driving instructor. Errors identified by the occupational therapist as contributing to an on-road Fail score were summed and ranked by frequency. The relationship between each error and Pass and Fail group was investigated using Fisher's Exact Test (for expected cell frequencies <5). The phi coefficient was computed as a measure of association of each error to the Pass and Fail rating.

8.2 Results for the Standard Assessment Group (n=60)

8.2.1 On-Road Assessment

Twenty-one of the 60 participants (35%) failed the on-road driving assessment, with no difference in Fail rates between males and females (13 of 36 males Failed, 8 of 24 females Failed; Fisher's Exact Test, two-tailed p = 1.0). The mean age of the Fail group was 80.7 years and 76.4 years for the Pass group (Mann-Whitney U test, two-tailed, z = -1.88, p = .06). Forty-eight participants started the on-road assessment from Burwood Hospital with the remaining 12 starting the assessment from home. There was no difference in Pass and Fail rates between those who started the assessment from Burwood hospital versus home (Mann-Whitney U test, two-tailed, z = -.54, p = .59). Table 8-2 summarizes the results of the Pass and Fail groups, including the number of drivers rated at each level of the driving scale score.

Table 8-2. Characteristics of on-road assessment Pass and Fail groups for n=60

<table>
<tr><th></th><th>Driving Scale score</th><th>Number of participants</th><th>Total</th><th>Sex</th><th>Age</th></tr>
<tr><td rowspan="6">Fail</td><td>0</td><td>0</td><td rowspan="6">n=21 Fail
(35.0% of sample)</td><td rowspan="3">n=8 females (33.3% of females)</td><td rowspan="6">Mean age
80.7 years</td></tr>
<tr><td>1</td><td>6</td></tr>
<tr><td>2</td><td>6</td></tr>
<tr><td>3</td><td>8</td><td rowspan="3">n=13 males (36.1% of males)</td></tr>
<tr><td>4</td><td>0</td></tr>
<tr><td>5</td><td>1</td></tr>
<tr><td rowspan="5">Pass</td><td>6</td><td>4</td><td rowspan="5">n=39 Pass
(65.0% of sample)</td><td rowspan="2">n=16 females (66.7% of females)</td><td rowspan="5">Mean age
76.4 years</td></tr>
<tr><td>7</td><td>12</td></tr>
<tr><td>8</td><td>14</td><td rowspan="3">n=23 males (63.9% of males)</td></tr>
<tr><td>9</td><td>9</td></tr>
<tr><td>10</td><td>0</td></tr>
</table>

In comparison to the spread of scores in the Healthy Older Drivers, study where the majority of scores clustered around a driving scale score of five or six on either side of the Pass/Fail divide, the driving scale scores in the current study were bimodal, with most scores clustering in the mid range of the Pass and Fail scores. This could suggest that referred drivers were more likely to be considered clean Pass and Fail outcomes rather than scoring in the borderline range. It could also show differences between ratings of the occupational therapist used in the Healthy Older Drivers study and the current study. Finally, it could also indicate a difference in driver behaviour as drivers in the Healthy Older Drivers study did not suffer consequences for poor performance which could have led to a more relaxed approach with a resultant clustering in the middle of the driving scale values range.

Eighty-six percent of test measures returning significant Shapiro-Wilk W scores, indicating that the majority of the data were non-normally distributed. These results are presented in Appendix Q.

8.2.2 Pass and Fail Groups: Significant Differences and Effect Sizes

Details of the ability of independent variables to discriminate between Pass and Fail groups are shown in Table 8-3 and Table 8-4 along with Cohen's effect sizes for parametric data and Cohen's effect-size for rank-transformed variables for non-parametric data (Hopkins, 2004).

Most *SMCTests*, with the exception of the Tracking and Arrows Perception tests, showed differences between Pass and Fail outcomes. Differences in the Ballistic Movement measures were in the direction that participants in the Fail group had slower reaction and movement times. In the Divided Attention test the Fail group correctly identified the direction of fewer arrows and omitted responses more frequently. In the Complex Attention test the Fail group had slower reaction and movement times and a higher number of invalid trials. In the Planning test the Fail group made more road position errors, remained in error for a greater amount of time, drove over more road hazards, and were involved in a greater number of crashes.

Table 8-3. Comparison of on-road Pass and Fail groups for non-normally distributed variables using Mann-Whitney *U* tests

Test Measure	Median Pass group (n=39)	Median Fail group (n=21)	Mann-Whitney *U* *p*-value (**p*<.05)	Cohen-type effect size[1]
Gender (1 = male, 2 = female)	1.0	1.0	.827	-0.06
Age (years)	79.0	80.0	.060	0.53
Handedness (1 = right, 2 = left)	1.0	1.0	.646	0.12
Ballistic Movement Test				
Reaction time, right hand (ms)	409	527	**.001***	1.00
Reaction time, left hand (ms)	409	462	**.001***	1.03

Continued on following page

				Continued from previous page
Reaction time, grand mean (ms)	408	548	**.000***	1.10
Movement time, right hand (ms)	270	361	**.001***	1.08
Movement time, left hand (ms)	270	350	**.000***	1.18
Movement time, grand mean (ms)	265	359	**.000***	1.18
Total reaction and movement times, right hand (ms)	748	878	**.000***	1.20
Total reaction and movement times, left hand (ms)	684	960	**.000***	1.20
Total reaction and movement times, grand mean (ms)	686	907	**.000***	1.29
Tracking Tests				
Sine tracking run 1 error (mm)	24.4	27.4	.190	0.36
Sine tracking run 2 error (mm)	14.5	18.9	.185	0.36
Random tracking run 1 error (mm)	12.9	15.5	.069	0.52
Random tracking run 2 error (mm)	11.7	15.9	.072	0.51
Arrows Perception Test				
Number of arrows correct	12.0	12.0	1.00	0.00
Omission of arrows response	0.0	0.0	.075	0.46
Divided Attention Test				
Tracking error (mm)	13.5	14.6	.137	0.37
Number of arrows correct	11.0	9.0	**.039***	-0.52
Omission of arrows response	0.0	2.0	**.006***	0.70
Complex Attention test				
Reaction time (ms)	682	903	**.003***	0.85
Movement time (ms)	463	486	.102	0.48
Movement time standard deviation (ms)	89	128	.083	0.48
Number of lapse errors	1.0	1.0	.366	0.24
Number of invalid trials	0.0	1.0	**.006***	0.75
Planning Test				
Lateral road position error (mm)	3.0	3.6	**.029***	0.71
Duration of positional faults (s)	11.6	20.2	**.023***	0.69
Distance travelled (m)	3.2	3.4	.553	0.18
Intersection safety margin (mm)	17.0	3.0	.369	-0.44
Number of hazards hit	2.0	3.0	**.015***	0.61
Number of crashes	2.0	4.0	**.036***	0.58

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value.

Table 8-4. Comparison of on-road Pass and Fail groups for normally distributed variables using *t*-tests

Test Measure	Mean Pass group (n=39)	Mean Fail group (n=21)	*t*-test *p* value (***p*<.05)	Cohen's *d* effect size[1]
Ballistic Movement Test				
Peak velocity, right hand (ms)	757	559	**.000***	-1.14
Peak velocity, left hand (ms)	743	555	**.000***	-1.01
Peak velocity, grand mean (ms)	750	557	**.000***	-1.09
Complex Attention test				
Total mean movement and reaction times (ms)	1205	1427	**.005***	0.79
Reaction time standard deviation (ms)	256	316	.051	0.55

[1]Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment.

8.2.3 Classification of On-Road Pass and Fail

Selection of variables offered to the BLR model followed the process described in Section 6.1.3.2. First, pragmatic decisions were made about the variables that would be excluded from the model. The list of the 11 excluded measures along with the reasons for exclusion can be found in Appendix R.

The next step was to examine the tolerance levels of the remaining variables using the SPSS function ‘collinearity diagnostics’. The effect sizes were ranked and the top 20 independent variables that were related to the dependent Pass/Fail outcome were checked for collinearity.

Five variables were deleted due to tolerance values <0.2 (see Section 6.1.3.2 for a description of tolerance testing and cutoff values). These variables were Ballistic Movement total time grand mean, Divided Attention omission of arrows response, Random Tracking run 1, Complex Attention total mean movement and reaction times, and Ballistic Movement movement time grand mean. Of the remaining 15 variables, the 12 variables with the highest effect size were offered to model (i.e., one variable per 5 participants). In order of highest to lowest effect size these variables were: Ballistic Movement reaction time grand mean, Ballistic Movement peak velocity grand mean, Complex Attention reaction time, Complex Attention number of invalid trials, Planning hazards hit, Planning duration of positional faults, Planning lateral road position error, Planning number of crashes, Divided Attention number of arrows correct, Complex Attention reaction time standard deviation, Age, and Random Tracking run 2.

These 12 measures were entered into the BLR using a step-wise backwards elimination procedure. The model accepted four measures: Ballistic Movement peak velocity grand mean, Complex Attention number of invalid trials, Complex Attention reaction time standard deviation, and Random Tracking run 2. These measures accounted for 47% of the variance in the on-road outcome (Nagelkerke R^2). The ROC AUC for the BLR model was .86 ($z = 6.55$; $p < .001$, 95% CI: .75–.94). The sensitivities, specificities, and total classification accuracies of the BLR across a range of cut-points are shown in Table 8-5.

As described in Section 6.1.3.6, two cut-points were compared for each model, with details included in Table 8-5. Using a default criterion value of 0.5 (‘Default’ cut-point), the model correctly classified 50 of 60 participants (83.3%) into on-road Pass or Fail groups with a negative predictive value of 83.7%, and positive predictive value of 82.4%.

Table 8-5. The sensitivities, specificities, and classification accuracies of the BLR model at different cut-points including the ‘Default’, and ‘Optimized’, cut-points

Criterion[1]	Sensitivity (%)	Specificity (%)	Mean Sensitivity & Specificity (%)	Accuracy (%)
≥ 0.10	95.2	33.3	64.3	55.0
≥ 0.20	90.5	59.0	74.7	70.0
≥ 0.30	85.7	69.2	77.5	75.0
≥ 0.39 ‘Optimized’	76.2	84.6	80.4	81.7
≥ 0.50 ‘Default’	66.7	92.3	79.5	83.3
≥ 0.60	52.4	92.3	72.3	78.3
≥ 0.70	38.1	94.9	66.5	75.0
≥ 0.80	23.8	97.4	60.6	71.7
≥ 0.90	19.1	97.4	58.2	70.0

[1]Criteria displayed are the predicted probabilities of receiving a Fail score.

The cut-point for the highest mean sensitivity and specificity value (mean = 80.4%, cut-point = 0.39, ‘Optimized’ cut-point) correctly classified 49 of 60 participants (81.7%) with a negative predictive value of 86.8%, and positive predictive value of 72.7%.

The 60 iterations generated by leave-one-out cross-validation reduced the accuracy of the ‘Default’ cut-point from 83.3% to 68.3% (41/60 correctly classified), sensitivity from 66.7% to 47.6%, specificity from 92.3% to 79.5%, negative predictive value from 83.7% to 73.8%, and positive predictive value from 82.4% to 55.6%. Leave-one-out cross-validation dropped the overall accuracy of the ‘Optimized’ cut-point from 81.7% to 63.3%, sensitivity from

76.2% to 52.4%, specificity from 84.6% to 69.2%, negative predictive value from 86.8% to 73.0%, and positive prediction value from 72.7% to 47.8%. For a summary of these results see Table 8-6.

Of the 60 iterations of the leave-one-out cross-validation, 16 contained a different set of measures from those used in the classification model: Ballistic Movement peak velocity grand mean, Complex Attention number of invalid trials, Complex Attention reaction time standard deviation, and Random Tracking run 2. Ballistic Movement peak velocity grand mean and Complex Attention invalid trials appeared in all 60 iterations, and were the sole tests used in 14 of the 16 iterations that accepted different tests from the classification model. For the remaining two iterations, Planning duration of positional faults was found in both and Planning hazards hit in just one.

8.3 Results for the Extended Assessment Group (n=32)

8.3.1 Diagnosis and Severity

Eight of the 32 participants were classified as MCI and the remaining 24 as Alzheimer's dementia. Of the eight with MCI, five met the criteria for Amnestic MCI multiple domain and three met the criteria for Amnestic MCI single domain (only memory significantly affected). Of those who met the criteria for Alzheimer's dementia, none fell in the SMMSE range of ≤9 indicative of severe impairment, six fell in the range of 10-20 indicative of moderate impairment, nine fell in the range of 21-24 indicative of mild impairment, and nine fell into the 'intact' range of 25+. Of those who fitted the criteria for MCI, one scored 19 on the SMMSE (in the moderate impairment range), one scored 24 (in the mild impairment range), and the rest scored between 25 and 28 (in the 'intact' range). Even if significant cognitive impairment was evident on testing, the defining criterion for separating MCI from Alzheimer's dementia was the informant report indicating the extent of functional impairment.

Table 8-6. The sensitivities, specificities and overall classification accuracies of the BLR model for the standard assessment sample at three different cut-points including the accuracy following leave-one-out cross-validation

Model	Classification					Following leave-one-out cross-validation				
	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)
BLR, n=60										
‘Default’ cut-point	66.7	92.3	83.7	82.4	83.3	47.6	79.5	73.8	55.6	68.3
‘Optimized’ cut-point	76.2	84.6	86.8	72.7	81.7	52.4	69.2	73.0	47.8	63.3

8.3.2 On-Road Assessment

Eleven of the 32 participants (34%) failed the on-road driving assessment with no difference in failure rates between males and females (9 males, 2 females; Fisher's Exact Test, two-tailed p = 0.43). The mean age of the Fail group was significantly higher than the Pass group at 79.2 years and 74.7 years respectively (t-test, two-tailed, t = -2.412, p = .02). Twenty-seven participants started the on-road assessment from Burwood hospital with the remaining five starting the assessment from home. There was no difference in Pass and Fail rates between those who started the assessment from Burwood versus home (Mann-Whitney U test, two-tailed, z = -.725, p = .47). There was a trend for drivers with dementia to be more likely to receive an on-road fail, but this was not significant (12.5% of the MCI group Failed versus 41.7% of the Alzheimer's group, Mann-Whitney U test, two-tailed, z = -1.481, p = .14). Table 8-7 summarizes the results of the Pass and Fail groups, including the number of drivers rated at each level of the driving scale score with the scores of the eight MCI participants noted.

Table 8-7. Characteristics of on-road assessment Pass and Fail groups for n=32

<table>
<tr><th></th><th>Driving Scale score</th><th>Number of participants</th><th>Total</th><th>Sex</th><th>Age</th></tr>
<tr><td rowspan="6">Fail</td><td>0</td><td>0</td><td rowspan="6">n=11 Fail (34.4% of sample)</td><td rowspan="3">n=2 females (20.0% of females)</td><td rowspan="6">Mean age 79.2 years</td></tr>
<tr><td>1</td><td>4</td></tr>
<tr><td>2</td><td>2 (1 MCI)</td></tr>
<tr><td>3</td><td>5</td><td rowspan="3">n=9 males (41.0% of males)</td></tr>
<tr><td>4</td><td>0</td></tr>
<tr><td>5</td><td>0</td></tr>
<tr><td rowspan="5">Pass</td><td>6</td><td>3 (3 MCI)</td><td rowspan="5">n=21 Pass (65.6% of sample)</td><td rowspan="2">n=8 females (80.0% of females)</td><td rowspan="5">Mean age 74.7 years</td></tr>
<tr><td>7</td><td>6 (1 MCI)</td></tr>
<tr><td>8</td><td>6 (1 MCI)</td><td rowspan="3">n=13 males (59.1% of males)</td></tr>
<tr><td>9</td><td>6 (2 MCI)</td></tr>
<tr><td>10</td><td>0</td></tr>
</table>

The driving scale scores for the n=32 were again bimodal, with most scores clustering in the mid range of the Pass and Fail scores. Possible reasons for this are the same as for those discussed in Section 8.2.1. A breakdown of on-road Pass and Fail outcome by MCI group and Alzheimer's dementia group broken down by severity is given in Table 8-8.

Table 8-8. Number and percentage of drivers with MCI and with Alzheimer's by dementia severity who Passed and Failed the on-road assessment

		Dementia severity		
On-road outcome	MCI	Intact	Mild	Moderate
Pass	7	6	6	2
Fail	1	3	3	4
Percentage of group who failed	12.5%	33.3%	33.3%	66.7%

8.3.3 Pass and Fail Groups: Significant Differences and Effect Sizes

Before the data were tested for normality, four measures were deleted from further analysis. No driver reported using a motorcycle as a form of transport so this item from the Driving Habits Questionnaire was removed. All 32 participants reported wearing a seat-belt while driving, and all participants reported driving both within their immediate neighbourhood and within the wider area of Christchurch city.

Normality testing showed that 81% of test measures returned significant Shapiro-Wilk W scores (i.e., were not normally distributed). These results are displayed in Appendix S.

There were no officially recorded crashes for any participant in the five years prior to completion of their on-road assessment.

Details of the ability for independent variables to discriminate between Pass and Fail groups are found in Table 8-9, Table 8-10, Table 8-11, and Table 8-12. Most of the *SMCTests* measures, excluding the Tracking tests and Arrows Perception, had at least one measure that was different between Pass and Fail groups. A small number of demographic and cognitive tests were different between Pass and Fail groups. Drivers in the Fail group had had their driving licences for a greater number of years than those in the Pass group, and were less likely to be driving at the time of the cognitive testing (i.e., their referrers had asked them to cease driving until they had completed the driving assessment). Drivers in the Fail group took longer to complete the Trail Making Test A, had higher error scores on the ADAS-Cog Total error score, and lower scaled scores on the Category Fluency test. The Ballistic Movement test had differences in the direction that participants in the Fail group had slower reaction and movement times. For the Divided Attention test the Fail group correctly identified the direction of fewer arrows, and omitted responses more frequently. For the Complex Attention test the Fail group had a larger reaction time standard deviation, a higher movement time standard deviation, and a higher number of invalid trials. For the Planning test the Fail group selected smaller gaps at intersections. Drivers in the Fail group reported driving to distant towns within the South Island of New Zealand in the previous 12 months more frequently than the Pass group.

The self-reported presence of medical conditions were compared between Pass and Fail groups using Fisher's Exact Test (see Table 8-13), while the number of conditions reported, medications taken for reported conditions, and amount of distress associated with medical conditions were compared between groups using Mann-Whitney *U* tests (these measures all had significant Shapiro-Wilk W scores indicating their data were non-normally distributed). As no participants in the n=32 sample reported stroke, Parkinson's disease, thyroid problems, multiple sclerosis, osteoporosis, head injury, diabetic retinopathy or retinal detachment these conditions do not appear in Table 8-13.

Having heart disease was associated with an increased likelihood of failing the on-road assessment (p = .003) with a Cohen-type ranked effect sizes of 1.24. None of the eye conditions reported were significantly associated with on-road outcome. There was no significant difference between on-road Pass and Fail groups in the number of medical conditions endorsed (z = -1.509, p = .131), frequency with which medications were taken for medical conditions (z = -0.292, p = .771), or the amount of distress associated with medical conditions (z = -0.722, p = .470). Although the Fail group had trends for higher mean ranked scores for both the number of medical conditions and number of medications taken, the Pass group had a trend for a higher ranked mean score for the amount of distress associated with medical conditions, which is in opposite to the direction that expected.

8.3.4 Classification of On-Road Pass and Fail

8.3.4.1 Binary Logistic Regression – extended assessment model 1

Selection of variables that were offered to the BLR model followed the process described in Section 6.1.3.2. First, pragmatic decisions were made about the variables that would be excluded from being offered to the model. A list of the 19 measures along with the reasons for exclusion can be found in Appendix T.

The next step was to examine the tolerance levels of the remaining variables using the SPSS function 'collinearity diagnostics'. The effect sizes were ranked and the top 15 independent variables that were related to the dependent Pass/Fail outcome were checked for collinearity.

Table 8-9. Comparison of on-road Pass and Fail groups for non-normally distributed demographic and driving variables using Mann-Whitney U tests.

Test Measure	Median Pass group (n=21)	Median Fail group (n=11)	Mann-Whitney *U p*-value (***p*<.05)	Cohen-type effect size[1]
Gender (1 = male, 2 = female)	1.0	1.0	.827	-0.44
Handedness (1 = right, 2 = left)	1.0	1.0	.646	0.16
Diagnosis (1 = MCI, 2 = Alzheimer's)	1.0	1.0	.139	0.60
Years of Education	11.0	12.0	.952	0.02
Years of Driving	51.0	62.0	**.029***	0.86
Demerit points earned in the previous 5 years	0.0	0.0	.413	0.32
Forms of transport used regularly (binary):				
drive own car	1.0	1.0	**.011***	-0.83
taken as passenger	1.0	1.0	.298	0.45
taxi	0.0	0.0	.969	-0.01
bicycle	0.0	0.0	.469	-0.31
bus	0.0	0.0	.866	0.06
walking	1.0	1.0	.939	0.03
Wear glasses when driving	1.0	1.0	.373	0.34
Self-rated quality of driving (ordinal)[2]	4.0	4.0	1.00	0.00
People who during the last year have driven:				
to neighbouring towns	1.0	1.0	.379	-0.31
to distant towns	0.0	1.0	**.003***	1.23
to places outside of the South Island	0.0	0.0	.469	-0.31

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value.[2]Self-rated driving was rated on an ordinal scale as follows: 1 = poor, 3 = average, 5 = excellent.

Table 8-10. Comparison of on-road Pass and Fail groups for non-normally distributed cognitive variables using Mann-Whitney *U* tests.

Test Measure	Median Pass group (n=21)	Median Fail group (n=11)	Mann-Whitney *U* *p*-value (***p<.05**)	Cohen-type effect size[1]
Geriatric Depression Scale	2.0	1.0	.088	-0.65
Mini-Mental State Exam	25.0	24.0	.133	-0.55
VOSP Incomplete Letters (z-score)	0.13	-0.66	.376	-0.33
Colour-Word Interference, Colour naming scaled score	6.0	4.0	.243	-0.42
Colour-Word Interference, Interference scaled score	6.0	1.0	.323	-0.22
Trail Making Test A (z-score)[2]	0.47	2.41	**.045***	0.80
Trail Making Test B (z-score)[2]	1.91	10.02	.068	0.68
Judgement of Line Orientation (percentile)	56.0	56.0	.572	-0.20
Rey Complex Figure copy (ordinal score)[3]	1.0	5.0	.133	0.59
Rey Complex Figure immediate recall (T score)	31.0	24.0	.796	-0.09
Activities of Daily Living Scales				
The Informant Questionnaire on Cognitive Decline in the Elderly	4.0	4.1	.952	-0.02

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value. [2]Higher z-scores for Trail Making Test A and B score equate to longer times required to complete the task relevant to the cognitively-unimpaired standardization group. [3]The Rey Complex Figure standardized data for the copy trial does not give scores across the spectrum of the distribution, but only up to the 16^{th} percentile. An ordinal score was assigned for the reported percentiles as such: 1 = >16^{th} percentile, 2 = 11–16^{th} percentile, 3 = 6–10^{th} percentile, 4 = 2–5^{th} percentile, 5 = ≤1^{st} percentile. Thus the lower the ordinal score, the more impaired the copy performance was.

Table 8-11. Comparison of on-road Pass and Fail groups for non-normally distributed *SMCTests* variables using Mann-Whitney *U* tests.

Test Measure	Median Pass group (n=21)	Median Fail group (n=11)	Mann-Whitney *U* *p*-value (***p<.05**)	Cohen-type effect size[1]
Ballistic Movement Test				
Reaction time, right hand (ms)	423	496	.77	0.71
Reaction time, left hand (ms)	415	444	**.005***	1.11
Reaction time, grand mean (ms)	408	548	**.015***	1.04

Continued on following page

Continued from previous page

Test Measure	Median Pass group (n=21)	Median Fail group (n=11)	Mann-Whitney *U* *p*-value (***p*<.05*)*	Cohen-type effect size[1]
Movement time, right hand (ms)	257	357	**.031***	0.98
Movement time, left hand (ms)	259	346	**.016***	1.04
Movement time, grand mean (ms)	259	342	**.013***	1.07
Total reaction and movement times, right hand (ms)	665	821	**.034***	0.95
Total reaction and movement times, left hand (ms)	683	1036	**.004***	1.31
Total reaction and movement times, grand mean (ms)	674	907	**.006***	1.23
Peak velocity, right hand (ms)	867	610	**.014***	-1.09
Tracking Tests				
Sine tracking run 1 error (mm)	22.0	26.5	.827	0.13
Sine tracking run 2 error (mm)	15.1	14.5	.677	0.08
Random tracking run 1 error (mm)	11.0	14.5	.592	0.16
Random tracking run 2 error (mm)	10.8	15.7	.416	0.23
Arrows Perception Test				
Number of arrows correct	11.0	11.0	.388	0.29
Omission of arrows response	0.0	0.0	.272	0.40
Divided Attention Test				
Number of arrows correct	11.0	11.0	**.036***	-0.84
Omission of arrows response	0.0	0.0	**.025***	0.87
Complex Attention test				
Reaction time standard deviation (ms)	190	351	**.045***	1.07
Movement time standard deviation (ms)	83	130	**.039***	0.86
Number of lapse errors	0.0	1.0	.635	0.19
Number of invalid trials	0.0	1.0	**.025***	0.86
Planning Test				
Lateral road position error (mm)	3.1	3.5	.310	0.38
Duration of positional faults (s)	8.9	20.2	.088	0.65
Distance travelled (m)	3.5	3.3	.473	-0.29
Intersection safety margin (mm)	23.0	0.0	**.046***	-0.81
Number of hazards hit	2.0	3.0	.268	0.45
Number of crashes	1.0	3.0	.195	0.48

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. Median scores do not necessarily represent the direction of the effect as this is based on a calculation that utilizes mean ranks rather than the median value.

Table 8-12. Comparison of on-road Pass and Fail groups for normally distributed variables using *t*-tests.

Test Measure	Mean Pass group (n=21)	Mean Fail group (n=11)	*t*-test *p* value (***p*<.05*)*	Cohen's *d* effect size[1]
Age	74.7	79.2	.077	0.59
Road code questions	10.4	9.5	.211	-0.48
Wechsler Test of Adult Reading	105.4	106.2	.846	0.07
ADAS-Cog Total score[2]	17.6	22.6	**.027***	0.87
VOSP Silhouettes (z-score)	-0.52	-0.51	.989	0.01
Colour-Word Interference, Word reading scaled score	8.0	5.5	.054	-0.75
Letter Fluency scaled score	8.8	6.0	.063	-0.73
Category Fluency scaled score	6.3	3.9	**.010***	-1.03
Letter-number Sequencing scaled score	8.5	7.0	.220	-0.48
Block Design scaled score	8.8	6.8	.130	-0.58
Activities of Daily Living Scales				
Four-Item Instrumental Activities of Daily Living Scale	7.1	6.6	.507	-0.25
Alzheimer's dementia Activities of Daily Living International Scale	1.2	1.0	.406	-0.31
Ballistic Movement Test				
Peak velocity, left hand (ms)	767	588	**.009***	-1.01
Peak velocity, grand mean (ms)	779	587	**.003***	-0.90
Divided Attention Test				
Tracking error (mm)	13.5	15.2	.429	0.30
Complex Attention test				
Reaction time (ms)	733	850	.112	0.61
Movement time (ms)	462	518	.271	0.42
Total mean movement and reaction times (ms)	1204	1384	.085	0.66

[1]Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment. [2]Higher ADAS-Cog Total scores indicate a higher number of errors performed.

Table 8-13. Comparison of reported medical conditions for on-road Pass and Fail groups and the Cohen-type effect size for ranked transformed variables

Medical Condition	% endorsed in Pass group (n=21)	% endorsed in Fail group (n=11)	Fisher's Exact Test (two-tailed, ***p*<.05**)	Cohen-type effect size[1]
Heart disease	4.7%	54.5%	**.003***	1.24
Cancer	9.5%	0%	.534	-0.45
High blood pressure	19.0%	18.2%	1.00	-0.02
High cholesterol	9.5%	36.4%	.148	0.65
Sleep apnoea	9.5%	27.3%	.310	0.45
Diabetes	0%	9.1%	.167	0.43
Arthritis	33.3%	45.5%	.703	0.24
Depression	23.8%	9.1%	.637	-0.39
Anxiety	14.3%	0%	.534	-0.56
Surgery	14.3%	18.2%	1.00	0.10
Falls	14.3%	9.1%	1.00	-0.16
Cataracts	28.6%	9.1%	.374	-0.50
Macular degeneration	0%	9.1%	.344	0.43
Glaucoma	4.8%	9.1%	1.00	0.16

[1]The Cohen's-type effect size is calculated using the mean ranks of Pass and Fail groups (Hopkins, 2004). Positive effect sizes show that a higher score on the measure was related to an increased likelihood of failing the on-road assessment, while negative effect sizes show a lower score on the measure was related to an increased likelihood of failing the on-road assessment.

Three independent variables were deleted due to tolerance values <0.2 (see Section 6.1.3.2 for a description of tolerance testing and cutoff values). The three deleted variables were Ballistic Movement total reaction and movement times grand mean, Ballistic Movement peak velocity grand mean, and Divided Attention number of arrows correct. Of the remaining 12 variables remained, the six variables with the highest effect size were offered to the model. In order of highest to lowest effect size these variables were: Heart disease, Ballistic Movement movement time grand mean, Ballistic Movement reaction time grand mean, Category Fluency scaled score, ADAS-Cog Total score, and Complex Attention number of invalid trials.

These six measures were entered into a BLR model using a step-wise backwards elimination procedure. The model accepted three measures: Heart disease, Category Fluency scaled score, and ADAS-Cog Total score. These measures accounted for 72% of the variance in the on-road outcome (Nagelkerke R^2). The ROC AUC for the BLR model was .92 (z = 7.114; p<.001, 95% CI: .78–.99). The sensitivities, specificities, and classification accuracies of the BLR across a range of cut-points are shown in Table 8-14.

As with the n=60 standard assessment sample, two cut-points were compared for each model. The 'Default' cut-point of 0.50 correctly classified 30 of 32 participants (93.8%) into on-road Pass or Fail groups with a negative predictive value of 95.2%, and positive predictive value of 90.9%. The 'Optimized' cut-point of 0.53 correctly classified 31 of 32 participants (96.9%) into on-road groups with a negative predictive value of 95.5%, and positive predictive value of 100.0%.

The 32 iterations generated by leave-one-out cross-validation reduced the accuracy of the 'Default' cut-point from 93.8% to 71.9%, sensitivity from 90.9% to 54.6%, specificity from 95.2% to 81.0%, negative predictive value from 95.2% to 60.0%, and positive predictive value from 90.9% to 60.0%. The accuracy of the 'Optimized' cut-point reduced from 96.9% to 71.9%, sensitivity from 90.9% to 54.6%, specificity from 100.0% to 81.0%, negative predictive value from 95.5% to 77.3%, and positive predictive value from 100.0% to 60.0%.

Table 8-14. The sensitivities, specificities and classification accuracies of the BLR model at different cut-points including the 'Default', and 'Optimized' cut-points

Criterion[1]	Sensitivity (%)	Specificity (%)	Mean Sensitivity & Specificity (%)	Accuracy (%)
≥ 0.10	90.9	71.4	81.2	78.1
≥ 0.20	90.9	76.2	83.6	81.3
≥ 0.30	90.9	81.0	85.9	84.4
≥ 0.40	90.9	95.2	93.1	93.8
≥ 0.50 'Default'	90.9	95.2	93.1	93.8
≥ 0.53 'Optimized'	90.9	100.0	95.5	96.9
≥ 0.60	72.7	100.0	86.4	90.6
≥ 0.70	72.7	100.0	86.4	90.6
≥ 0.80	54.6	100.0	77.3	84.4
≥ 0.90	54.6	100.0	77.3	84.4

[1]Criteria are the predicted probabilities of receiving a Fail score.

Note that the values following leave-one-out cross-validation are identical for both the Default and Optimized cut-point due to the closeness of these two points (0.50 and 0.53 respectively). For a summary of these results see .Table 8-15.

Of the 32 iterations of the leave-one-out cross-validation, 12 contained a different set of measures from the model formed by heart disease, Category Fluency scaled score, and ADAS-Cog Total score that were utilized in the classification. Ten of these 12 different models, however, did not include any additional tests from the six tests offered to the model, but rather dropped one of the three tests from the classification model. Eight of these models retained Heart disease and ADAS-Cog Total score while dropping Category Fluency scaled score, and a further two models retained Heart disease and Category Fluency scaled score and dropped ADAS-Cog Total score. Only two models contained a different test and in both instances this extra test was Ballistic Movement reaction time grand mean.

8.3.4.2 Binary Logistic Regression – extended assessment model 2

As described in Section 8.1.4.2, an alternate method for choosing variables to be offered to the BLR model was investigated. Instead of offering the six independent variables with the largest effect size between Pass and Fail groups to the model, the top two *SMCTests* measures (no more than one measure from a subtest), the top two cognitive measures (again no more than one measure from each test) and the top two medical illnesses were offered to the BLR model using the backwards step-wise selection method. The process prior to this selection was identical to the standard BLR reported above.

The six variables that met criteria outlined above were heart disease, high cholesterol, Ballistic Movement movement time grand mean, Complex Attention number of invalid trials, Category Fluency scaled score, and ADAS-Cog Total score. Five of these variables were the same as five used in the standard BLR model except for the exclusion of Ballistic Movement reaction time grand mean (this was bypassed due to the Ballistic Movement movement time grand mean measure already having been offered to the model) and the inclusion of high cholesterol.

Step-wise backwards elimination procedure resulted in a model with the same three measures as the n=32 extended assessment model 1: heart disease, Category Fluency scaled score, ADAS-Cog Total score. Because the same measures were accepted, model 2 was the same as model 1 in every respect including classification accuracy statistics and ROC AUC (see Section 8.3.4.1 and Table 8-14).

Table 8-15. The sensitivities, specificities and overall classification accuracies of the BLR model for the extended assessment model 1 at three different cut-points including the accuracy following leave-one-out cross-validation

Model	Classification					Following leave-one-out cross-validation				
	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)
BLR, n=32 model 1										
‘Default’ cut-point	90.9	95.2	95.2	90.9	93.8	54.6	81.0	77.3	60.0	71.9
‘Optimized’ cut-point	90.9	100.0	95.5	100.0	96.9	54.6	81.0	77.3	60.0	71.9

Although the model 2 classification model was the same as model 1, the leave-one-out cross-validation results could have been different since the six measures offered to the classification model were also offered to each iteration of the leave-one-out cross-validation, meaning that the deletion of Ballistic Movement reaction time grand mean and the inclusion of high cholesterol could have changed one or more iterations. The 32 iterations of leave-one-out cross-validation did indeed lead to slight changes compared to the extended assessment model 1. The accuracy of both the 'Default' and Optimized cut-points reduced 75.0%, sensitivity to 54.6%, specificity to 85.7%, negative predictive value to 78.3%, and positive predictive value to 66.7%. For a summary of these results see Table 8-16.

Twelve of the 32 iterations of leave-one-out cross-validation contained a different set of measures from the model formed by heart disease, Category Fluency scaled score, and ADAS-Cog Total score that were utilized in the classification. However, 11 of these 12 different models did not include any additional tests from the six tests offered to the model, but rather dropped one of the three tests from the classification model. Nine of these models retained Heart disease and ADAS-Cog Total score while dropping Category Fluency scaled score, and the remaining two models retained Heart disease and Category Fluency scaled score and dropped ADAS-Cog Total score. Only one model contained additional tests and in that case all six of the offered variables were accepted into the model.

To further clarify the relationship between the two extended assessment models, they are in practice the same model as they both include the same three independent measures. Extended assessment model 2 shows that by offering different measures to the model in a leave-one-out cross-validation can change the estimated predictive accuracies. The practical application of the extended assessment model is that only three measures would be performed and included in the BLR model and the model will subsequently be referred to in singular form.

8.3.4.3 Binary Logistic Regression – SMCTests Measures Only

As the n=32 extended assessment model appeared to do a better job at classifying on-road outcome than the n=60 standard assessment model, and as no *SMCTests* measures were accepted into the n=32 model, we decided to do an additional analysis to see the effect of offering only the measures that were available to the n=60 standard assessment sample to the n=32 sample: *SMCTests* measures, age, and gender. We were interested in whether the accuracy of this classification model would resemble the n=60 full sample model.

Table 8-16. The sensitivities, specificities and overall classification accuracies of the BLR model for the extended assessment model 2 at two different cut-points including the accuracy following leave-one-out cross-validation

Model	Classification					Following leave-one-out cross-validation				
	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)
BLR, n=32 model 2										
'Default' cut-point	90.9	95.2	95.2	90.9	93.8	54.6	85.7	78.3	66.7	75.0
'Optimized' cut-point	90.9	100.0	95.5	100.0	96.9	54.6	85.7	78.3	66.7	75.0

The 15 variables with the highest effect sizes were entered into 'collinearity diagnostics'. These variables include age and gender along with 13 *SMCTests* measures. Four variables were deleted due to tolerance values <0.2 (see Section 6.1.3.2 for a description of tolerance testing and cutoff values). The four deleted variables were Ballistic Movement reaction time grand mean, Ballistic Movement peak velocity grand mean, Ballistic Movement total time grand mean, and Divided Attention number of arrows correct. Of the remaining 11 variables, the six variables with the highest effect size were offered to the model. In order of highest to lowest effect size these variables were: Ballistic Movement movement time grand mean, Complex Attention number of invalid trials, Divided Attention omission of arrows response, Complex Attention reaction time standard deviation, Complex Attention movement time standard deviation, and Planning intersection safety margin.

These six measures were entered into a BLR model using a step-wise backwards elimination procedure. The model accepted two measures: Divided Attention omission of arrows response, and Planning intersection safety margin. These measures accounted for 37% of the variance in the on-road outcome (Nagelkerke R^2). The ROC AUC for the BLR model was .79 (z = 3.147; p=.002, 95% CI: .61–.91).

As the n=60 standard assessment model had a Nagelkerke R^2 value of 47% and an AUC of .86, the n=32 *SMCTests*-only model is clearly not as accurate at classification. As with previous models, two cut-points were investigated and leave-one-out cross-validation was performed and the results were very poor. Following leave-one-out cross-validation, the sensitivities of both cut-points were 27.3%, with positive predictive values also 27.3% indicating that this model would detect few true Fails and would misclassify many Passes as Fails. It appears that *SMCTests* measures were not useful in classifying on-road outcome in this sample.

8.4 On-Road Errors and On-Road Outcome

Driving errors conducted during the on-road assessment were recorded using an on-road error list described in Section 5.10.1. Both assessors were asked to note the occurrence of errors, whether they continued despite verbal instruction, and whether the assessor considered the error to be a contributing factor for an on-road Fail score. It was subsequently decided that the rating of whether the participant continued performing an error after instruction was not consistent between participants. This was because the chance to prove corrected behaviour might not occur during the remainder of the drive. Thus, these data were not further investigated. Table 8-17 displays the frequency that each error on the driving error checklist was rated by the two raters as occurring, and also Cronbach's α which determines the level of agreement between raters. There was a high degree of agreement between the raters with 11 error ratings with α values of .80 or above. The error 'Incorrect indication at an intersection' had a negative α of -0.11. This is due to no agreement between the two ratings for any of the 7 participants who had this error noted.

The two most frequently rated reasons for a Fail score were 'Decreased awareness of other road users' and 'Decreased awareness of environment', with the former reason being a reason for all 21 participants who failed the assessment (see Table 8-18). Five errors did not contribute to a Fail score for any participants. 'Incorrect indication at a roundabout' which had the highest occurrence in the sample as a whole (42 participants according to the occupational therapist's rating) was only considered by the occupational therapist to have directly contributed to one person's Fail score.

Eighteen errors were related to Pass and Fail outcomes (Table 8-18). All but four variables were in the expected direction of occurring more frequently in the Fail group. Three variables did not occur in the Fail group at all: 'Turned into incorrect lane on multi-lane road', 'Fails to give way to pedestrians at intersection', and 'Inappropriate use of arrow traffic lights'. The error 'Fails to follow pedestrian crossing rules' occurred once in the Fail group and twice in the Pass group.

The errors most related to a Fail outcome are shown by the largest phi coefficients and are similar to the occupational therapist's self-rated contribution of error to Fail scores (Table 8-18). Of the top 10 errors rated as contributing to Fail scores by the occupational therapist, 8 were ranked in the top ten by phi coefficient for errors associated with Fail scores: 'Decreased awareness of other road users', 'Decreased awareness of environment', 'Lack of scanning techniques', 'Inappropriate gap selection', 'Incorrect use of give way rules at intersections', 'Didn't react in time to situation / incorrect action taken', 'Driving too close to (or over) left line', and 'Driving above the speed limit'.

Table 8-17. Cronbach's α and 95% confidence interval for agreement on the commission of errors between raters

Type of Error	Frequency rated by occupational therapist	Frequency rated by driving instructor	Cronbach's α	95% confidence interval
Fails to follow pedestrian crossing rules	3	3	1.00	1.00 – 1.00
Decreased awareness of environment	21	20	0.98	.97 - .99
Decreased awareness of other road users	21	19	0.96	.94 - .98
Driving/starting in wrong gear	7	6	0.96	.93 - .97
Inappropriate gap selection	18	15	0.94	.98 - .96
Incorrect use of give way rules at intersections	14	13	0.92	.87 - .95
Driving below the speed limit	16	16	0.91	.84 - .94
Driving above the speed limit	29	27	0.89	.82 - .93
Incorrect indication at a roundabout	42	44	0.86	.77 - .92
Turned into incorrect lane on multi-lane road	4	7	0.83	.72 - .90
Inappropriate use of arrow traffic lights	2	1	0.80	.66 - .88
Driving too close to (or over) centre line	9	11	0.78	.64 - .87
Lack of scanning techniques	19	21	0.77	.62 - .86
Driving too close to (or over) left line	15	11	0.76	.60 - .86
Didn't react in time to situation / Incorrect action taken	13	8	0.76	.60 - .86
Lack of mirror use	21	24	0.76	.60 - .86
Problem cornering: speed or position	8	8	0.72	.54 - .84
Incorrect use of lanes in roundabout	3	4	0.71	.51 - .83
Fails to observe signs	8	9	0.69	.47 - .81
Immediate fail error (e.g. crash)	2	7	0.61	.34 - .77
Lack of blind spot check	17	23	0.59	.31 - .75
Following other cars too closely	12	3	0.56	.26 - .74
Didn't apply 12 second search	6	17	0.55	.24 - .73
Gear grinding/over-revving	5	1	0.50	.16 - .70
Stopping too closely behind cars	5	2	0.41	.01 - .65
Approaching intersections at excessive speed	2	6	0.36	-.07 - .62
Incorrect indication at an intersection	2	5	-0.11	-.86 - .34
Incorrect indication for lane changes	1	0	N/A	N/A
Fails to give way to pedestrians at intersection	1	0	N/A	N/A

Table 8-18. Frequency of errors rated as contributing to a Fail outcome by the occupational therapist, and the statistical relationship of errors to Pass and Fail outcomes

Error	Frequency rated as contributing to the Fail score	Fisher's Exact Test *p* value	Phi coefficient
Decreased awareness of other road users	21	<.01	1.00
Decreased awareness of environment	20	<.01	0.93
Lack of scanning techniques	15	<.01	0.78
Inappropriate gap selection	13	<.01	0.74
Driving above the speed limit	13	<.01	0.48
Incorrect use of give way rules at intersections	12	<.01	0.67
Didn't react in time to situation / incorrect action taken	10	<.01	0.55
Driving below the speed limit	8	0.01	0.35
Driving too close to (or over) centre line	7	<.01	0.47
Driving too close to (or over) left line	6	<.01	0.54
Problem cornering: speed or position	6	<.01	0.53
Lack of mirror use	6	<.01	0.49
Lack of blind spot check	5	<.01	0.39
Following other cars too closely	5	0.09	0.24
Gear grinding/over-revving	3	<.01	0.41
Didn't apply 12 second search	3	0.02	0.34
Fails to observe signs	3	0.02	0.33
Incorrect use of lanes in roundabout	3	0.04	0.31
Stopping too closely behind cars	3	0.05	0.28
Driving/starting in wrong gear	2	0.23	0.17
Incorrect indication at an intersection	1	0.12	0.25
Incorrect indication at a roundabout	1	0.56	0.10
Fails to follow pedestrian crossing rules	1	1.00	-0.01
Immediate fail error (e.g. crash)	0	0.10	0.25
Incorrect indication for lane changes	0	0.35	0.18
Fails to give way to pedestrians at intersection	0	1.00	-0.10
Inappropriate use of arrow traffic lights	0	0.54	-0.14
Turned into incorrect lane on multi-lane road	0	0.29	-0.20

As part of the Driving Error List, the assessors were asked to rate each participant's level of insight, awareness of their driving problems, receptiveness to feedback about their driving, and likelihood of benefiting from lessons. Only one person was rated as likely to benefit from lessons so this rating was not further explored. Ratings by the occupational therapist for the remaining three categories are given in Table 8-19. As is clear from the table, there was a pattern for those who passed the assessment to be rated more often as having intact insight, having an awareness of driving problems, and being more receptive to feedback following the drive. The distribution of these scores were significantly different (χ^2 = 41.83, p = <.001, χ^2 = 40.52, p = <.001, χ^2 = 21.21, p = <.001 respectively).

Table 8-19. Ratings for insight, awareness of driving problems, and receptiveness to feedback for drivers who passed and failed the on-road assessment

	On-road outcome	
Rating	Pass (number)	Fail (number)
Insight		
Yes	33	1
Limited	5	4
No	1	16
Awareness of driving problems		
Yes	29	1
Limited	7	1
No	3	19
Receptiveness to feedback		
Yes	38	10
Limited	0	1
No	1	10

It is not surprising that those in the Pass group were more often receptive to feedback as this would have included the recommendation that they could continue driving. However, around half of those who were told they could no longer drive were also rated as receptive to this feedback.

8.5 Discussion

8.5.1 Associations Between Independent Variables and On-Road Outcome

Many *SMCTests* measures and standard cognitive measures discriminated between on-road Pass and Fail groups with large effect sizes (<0.80). Many effect sizes for *SMCTests* were over 1.0, with the Ballistic Movement measures showing consistently large effects (.71–1.31). Several scores from the Complex Attention test also produced high effect sizes including reaction time (0.85), number of invalid trials (0.75), and total mean movement and reaction times (0.79). This contrasts with the Healthy Older Drivers study in which the two highest effect sizes were 0.80 for Trail Making Test B and 0.71 for Random Tracking run 1, with the majority of variables falling within the small to medium range (0.20 – 0.50). The effects sizes found in the current study are higher than those found in Reger et al.'s (2004) meta-analysis for the ability of cognitive tests to discriminate between drivers with dementia with positive or negative driving outcomes (determined for most studies following on-road assessment). The cognitive domain with the highest effect size in the Reger et al. study was visuospatial skills with an effect of 0.29.

Examining the effect sizes of cognitive measures in the n=32 subsample, the measure with the highest effect size (-1.03) was the Category Fluency scaled score which is a measure of generative speech that is reflective of executive function skills. The ADAS-Cog Total score which represents a comprehensive screen of the cognitive impairments found in Alzheimer's dementia also had a large effect size (0.87). The Trail Making Test A's effect size of 0.80 was larger than the effect size of 0.68 for Trail Making Test B. Part B of the test is more difficult to complete than part A and many participants (both Passes and Fails) performed poorly or could not finish part B, therefore reducing its discriminative ability for predicting on-road Pass and Fail outcomes. That an essentially simplistic test such as Trail Making Test A, which consists of drawing lines between circled numbers, much like a connect-the-dots drawing, can discriminate reasonably well between on-road Pass and Fail groups suggests that many participants were impaired to the point where more basic skills such as visual scanning and/or motor speed were impaired.

Contrary to our expectations, informant reports of functional impairments were not useful in deciding which drivers with cognitive impairment would Pass or Fail the on-road assessment. The trend of the Alzheimer's dementia Activities of Daily Living International Scale was in the opposite from expected direction in that drivers who failed the on-road assessment were rated as less impaired then drivers who passed. Although the trends for Four-Item Instrumental Activities of Daily Living Scale and the Alzheimer's dementia Activities of Daily Living International Scale were both in the expected direction that the fail group had poorer scores, the effect sizes were small (-0.25 and -0.31 respectively). Perhaps the information received from informants was biased, in that difficulties were minimized either unconsciously or on purpose. Alternatively, perhaps these activities included in the questionnaires do not correlate well with the driving task.

Of note in the extended assessment results is the strong effect size (1.23) between self-reports of participants driving their cars to distant towns in the South Island of New Zealand within the last 12 months. The effect was in the direction that the people who reported that they had driven long distances were more likely to *fail* the on-road assessment. This seems counter-intuitive, as we would expect that people driving further afield would be more confident in their driving abilities, and would hope that this confidence would be based at least to some extent on their actual level of driving ability. The fact that the relationship was in the opposite direction perhaps indicates that these people driving longer distances had poorer insight into their driving abilities and were placing themselves at increased risk of negative driving outcomes through their behaviour. The sex distribution of people who reported driving to distant towns was investigated, with 82% of respondents being male compared to 69% of the total sample. Perhaps there is a sex influence, whereby men are more likely to continue driving longer distances even in the face of poorer driving skills. This measure was not offered to the extended assessment model as it was eliminated at the first step for pragmatic reasons along with the other driving behaviour questions. This self-report measure would be worth investigating in future studies of drivers with dementia to confirm that the result was indeed spurious.

Another interesting outcome for self-reported driving behaviour was that drivers who reported that they were not currently driving their own car at the time of the cognitive testing were more likely to fail the on-road assessment (Cohen-type effect size of -0.83). Participants in this group had ceased driving for a variety of reasons. Some had had a recent physical health problem which had required driving cessation for a few months and whose ability to restart driving was then queried on cognitive grounds which lead to the referral for assessment. Some drivers had been told by their doctors at the time of referral that they should not drive until the outcome of their on-road assessment was known. However, not all participants who were told not to drive were complying with this instruction, and several confided to the experimenter that they had continued to drive, at least on one or two occasions due to having no other easy method of transportation. These people were noted as still driving their car regardless of what they were advised to do by their doctor. It is also possible that participants told the examiner that there were not driving when they actually were for fear of consequences for not following their doctor's orders. So the validity of this measure is questionable, but could suggest that those drivers with MCI or Alzheimer's referred for a driving assessment may be more likely to fail if they were not currently driving.

There were trends for many medical conditions to be more prevalent in the Pass group, including cancer, depression, anxiety, falls, and cataracts. Some potential problems with the way details regarding medical conditions were collected is mentioned in Section 8.5.2. In summary, medical conditions were self-reported often with a family member present who was able to confirm the diagnosis or prompt the participant to report ones they had forgotten. Hence, the results should be considered reliable, although the relationship between physician-reported medical conditions and on-road driving ability would be a useful future area of enquiry. The presence of heart disease discriminated between Pass and Fail groups with a large effect size (1.24) which indicates its potential usefulness for detecting those who might be more likely to Fail an on-road assessment. Interestingly, there was a trend for drivers who reported more distress associated with medical conditions to be more likely to *pass* the on-road assessment, which is in the opposite direction from that expected. Perhaps driving problems linked with distress associated with medical conditions needs a larger sample of driving behaviour to detect than provided by a 45-minute on-road driving assessment. Alternatively, perhaps those who reported distress mad more insight generally and were thus more likely to Pass the on-road assessment.

Interestingly, the 15-item road rules knowledge test failed to discriminate between on-road Pass and Fail groups. Also, many of the participants required assistance in understanding the intention of the questions and took a very long time to complete this 'short' test. This indicates that road rules tests may be too difficult for drivers with MCI and Alzheimer's dementia to complete (both Passes and Fails) and that road rules knowledge should instead be assessed in context as part of an on-road driving assessment.

Overall, the effects of most independent variables were in the expected direction that drivers who failed the on-road driving assessment on average had poorer scores. As mentioned in Section 5.1.3.2, Babyak (2004) has many objections to researchers constructing models on the basis of relationships between independent and dependent variables found within the same sample and suggests instead that researchers have clear ideas about the variables they want to investigate and to enter these variables into models on the basis of theory rather than association within the sample. The data provided in Table 8-3, **Error! Reference source not found.**, Table 8-9, Table 8-10, Table 8-11, and Table 8-12 include effect size differences between Pass and Fail groups and could provide fruitful measures to include in a model built with a new participant sample.

8.5.2 Classification of On-Road Driving Ability

Table 8-20 summarizes the accuracy statistics for the two n=32 models and the n=60 model which will be discussed separately below.

8.5.2.1 Standard Assessment Model (n=60)

The standard assessment classification model had lower sensitivities than the extended assessment model. The model could only be offered *SMCTests* measures and chronological age and it selected Ballistic Movement peak velocity grand mean, Complex Attention number of invalid trials, Complex Attention reaction time standard deviation, and Random Tracking run 1. Given the four measures accepted into the classification model, off-road testing would take around 30 min with the Ballistic Movement test, Complex Attention test and Sine and Random Tracking runs 1 and 2 needing to be performed using the *CanDAT* semi-portable system.

The leave-one-out cross-validation produced expected decrements in both sensitivity and specificity, although the drops in accuracy were not as large as the drops observed in the extended assessment model, which is likely due to the larger sample size of this model which allowed for a more robust classification model to be built. During the leave-one-out cross-validation, there were 16 iterations with models different from that of the classification model although only two of these 16 contained different measures, with Planning duration of positional faults making two appearances and Planning number of hazards hit making one appearance. Following leave-one-out cross-validation, the standard assessment model had negative predictive values in the 73.0% to 73.8% range for the two presented cut-points which indicates that around 27% of those with a predicted Pass are likely to actually receive a Fail on the on-road assessment. The positive predictive values range from 47.8% to 55.6% for the two cut-points which indicates that around 50% of participants who receive a Fail score on the off-road assessment would also be expected to receive a Fail score on the on-road assessment.

Table 8-20. The sensitivities, specificities and overall classification accuracies of the BLR models at three different cut-points for classification models along with their accuracy following leave-one-out cross-validation

Model	Classification					Following leave-one-out cross-validation				
	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Positive predictive value (%)	Accuracy (%)
BLR, n=60										
‘Default’ cut-point	66.7	92.3	83.7	82.4	83.3	47.6	79.5	73.8	55.6	68.3
‘Optimized’ cut-point	76.2	84.6	86.8	72.7	81.7	52.4	69.2	73.0	47.8	63.3
BLR, n=32 model 1										
‘Default’ cut-point	90.9	95.2	95.2	90.9	93.8	54.6	81.0	77.3	60.0	71.9
‘Optimized’ cut-point	90.9	100.0	95.5	100.0	96.9	54.6	81.0	77.3	60.0	71.9
BLR, n=32 model 2										
‘Default’ cut-point	90.9	95.2	95.2	90.9	93.8	54.6	85.7	78.3	66.7	75.0
‘Optimized’ cut-point	90.9	100.0	95.5	100.0	96.9	54.6	85.7	78.3	66.7	75.0

8.5.2.2 Extended Assessment Models (n=32)

The measures accepted into the extended assessment model were (i) the binary incidence of self-reported heart disease (with an incidence of 54.5% in the Fail group and 4.7% in the Pass group); (ii) Category Fluency scaled score, which is a measure of generative speech often used to measure executive function deficits; and (iii) the ADAS-Cog Total score, which is a reasonably comprehensive screen for cognitive impairments associated with Alzheimer's dementia. None of the *SMCTests* measures offered to the models were accepted into the model.

The most intriguing of the accepted measures is heart disease. Given that all participants met criteria for either a diagnosis of Alzheimer's dementia or MCI, the potential effects of heart disease on driving must be over and above the cognitive deficits measured by the ADAS-Cog and Category Fluency test. It is possible that heart disease measures some useful interaction of cognitive and sensory-motor measures that are not explained better by the measurement of these deficits on their own. It could also be that the method of collecting the heart disease information was faulty. Information was not sought from participants' doctors but rather from self-report. This was often done with a family member present which would reduce the chances that large numbers of participants were giving inaccurate information. Also, the reported absence of heart disease does not mean that vascular pathology does not exist but rather that symptoms may not have been apparent or reported. The finding of the importance of heart disease requires replication, in particular to see if reports from a participant's doctor are associated with on-road driving. If this association is confirmed, then the presence of heart disease in patients with dementia would be a valuable indicator for doctors to determine the possible driving risk profile of their patients. Given the three measures accepted into the classification model, off-road testing could take around 35-50 minutes to perform (ADAS-Cog approx 30-40 minutes for those with dementia plus 5-10 minutes for Category Fluency depending on whether its forerunner test Verbal Fluency was performed first).

During the leave-one-out cross-validation of the two extended assessment models there was a large decline in sensitivity, from 90.9% for both classification cut-points to 54.6%. Specificity was less affected but also dropped from the high 90s to the low to mid 80s. Following the leave-one-out cross-validation, the negative predictive values in the high 70s for both cut-points indicate that around 22% of those predicted to Pass the on-road road assessment would actually receive a Fail score. The positive predictive values of 60% to

66.7% indicate that around 37% of drivers predicted to Fail the on-road assessment would actually Pass.

In addition, a model formed using just *SMCTests* measures performed very poorly compared to the cognitive tests and medical condition data accepted into the model. This adds further evidence that *SMCTests* measures are of little value in helping classify on-road assessment outcomes in people with MCI and Alzheimer's dementia.

Overall, the extended assessment model produced higher sensitivities in classification than the standard assessment model but these lowered substantially following leave-one-out classification to be only slightly higher than the standard assessment model. This may indicate overfitting occurring in the extended assessment model, most likely due to the smaller sample size which would be expected to produce a less stable model than the standard assessment model. Age was not controlled for in any of the models, although many of the cognitive test scores were adjusted for age. Age did not have a high enough effect size to be offered to the extended assessment model, and although it made the top 12 of variables offered to the standard assessment model, it was not accepted by the classification model or any of the 60 iterations of the leave-one-out cross-validation. In the standard assessment sample, age discriminated Pass and Fail groups with a moderate effect size of 0.53. When investigating samples with restricted age ranges, such as in the current study and the Healthy Older Drivers study, it is less likely that age would appear as a predictor of driving ability than when a sample also includes middle-aged and younger drivers.

8.5.3 Interpretation of On-Road Error Results

Inter-rater agreement on the commission of driving errors was high for many items. This indicates that many of the listed errors were identified independently by each rater. Discrepancies in some ratings could be due to varying thresholds for each rater in defining when an error occurred, and also to different levels of attention to driving behaviour; the driving instructor had the dual task of guiding the drive as well as maintaining safety of the vehicle, whilst the occupational therapist was able to concentrate fully on driving behaviour. Reliability of the detection of errors is an important first step in determining the usefulness of a behavioural rating system.

The two measures most commonly cited as contributing to a Fail outcome were not errors per se, but rather a subjective judgment made by the occupational therapist and/or driving instructor that a driver exhibited decreased awareness of the environment and/or other road users. The fact that these two awareness measures were rated for all but one fail outcome suggests that a single measure of decreased awareness could be used instead of two separate ones. This result may indicate that subjective statements presumably based on both the commission of errors and an interpretation of general demeanor and behaviour during the assessment are useful for making decisions about driving ability. There was a high level of agreement between the subjective ratings of the occupational therapist as to which errors were especially important to a Fail outcome and the degree to which the frequency of these errors (whether or not they were rated as important) was associated with on-road Fail scores.

The current study found a number of potentially useful error measures that were related to on-road Pass and Fail outcomes and were also rated as useful contributors to identifying people with on-road Fail outcomes by an occupational therapist. Including an evidence-based error list as part of a non-standardized driving assessment would allow for systematic collection of error information that may contribute to more reliable Pass and Fail on-road outcomes.

8.5.4 Generalization of the Results

The participants in the current study were referrals from a driving assessment service that included both those referred by general practitioners and those referred from a memory assessment clinic in the same district. The exclusion criteria required that participants had not had a previous stroke or severe head injury in an attempt to recruit a sample of people whose cognitive impairments were due to MCI or Alzheimer’s dementia. Thus, the results cannot necessarily be expected to generalize well to those with histories of brain insult not due solely to a dementing process. Results may also not generalize well to a group of drivers with forms of dementia other than Alzheimer’s dementia.

The sample of 32 participants who agreed to additional testing may have been different from the remaining 28 participants in the full sample. As reported in Section 8.1.1, there was no difference in male/female composition of the extended and standard assessment groups but age almost reached significance (p = .051) for the standard assessment group being older (mean = 79.9 years in standard group versus 76.2 years in extended assessment group). Further unmeasured differences between groups may influence the generalizability of the results of the cognitive testing but will not affect the results of the *SMCTests* measures in the standard assessment sample.

Because of potential limitations in the recruiting process, the predictive value assigned to the individual tests used in the BLR models could be lower or higher for the general population of drivers with dementia.

8.5.5 Errors in Prediction of On-Road Assessment Outcome

As with the Healthy Older Drivers study, it is likely that some measures potentially useful for the prediction of driving ability were not included in the current study. A rating of dementia severity using the Clinical Dementia Rating may have been useful in the current study. As shown in Table 8-8, there was a trend for on-road failure rates to increase from MCI to Alzheimer's ratings of 'intact' and mild impairment and then to moderate impairment, but a larger sample than 32 would be required to raise confidence in these findings.

As with the Healthy Older Drivers study, the reliability and validity of the on-road driving assessment is unknown. Section 3.8 details potential problems with the assessment. It is important to note that the validity of any on-road test purporting to determine driver safety in dementia groups will likely never be known. Establishing predictive validity for the ability of the on-road assessment to predict subsequent negative driving outcomes such as crashes and traffic offences would be dependent on allowing all drivers with dementia to continue driving regardless of on-road assessment performance in order to determine the subsequent rate of negative driving events between Pass and Fail groups. Since drivers with dementia are a high-risk group for negative driving events, it in unlikely that this sort of study will ever be performed and that pragmatic decisions made by specialists will continue to inform which people with dementia are judged to be safe or unsafe on the road.

8.5.6 Comparison with Predictive Models of Driving in the Literature

As with the Healthy Older Drivers study, there are few studies in the driving literature with which to compare the results of the Dementia and Driving models. This is because few researchers have recruited an MCI and/or dementia sample and compared off-road testing to blinded on-road outcome. Two studies were found that constructed a classification model for on-road performance that included a control group mixed with the dementia group (Hunt et al., 1993; Fitten et al., 1995). Such models are likely to be biased in reporting specificity since a control would be expected to perform better on both the off- and on-road testing. This in turn would increase the overall accuracy of the model, even if the sensitivity for classifying those who failed (more likely to be those with dementia) was moderate to low in size, particularly if there were more controls recruited than people with dementia. Hence, only studies that compare off- and on-road testing within a sample of people with MCI or dementia were compared to the models in the current study. Four studies were found that met these criteria (see Table 8-21).

Snellgrove (2000) recruited 115 community dwelling adults with MCI (CDR score of 0.5, n=23) and dementia (CDR score of 1.0, n=92, 65% Alzheimers dementia, 30% vascular dementia, 13% frontotemporal dementia, 2% Lewy-Body dementia) and used a simple maze completion task to classify Pass and Fail ratings for an on-road driving assessment using BLR. Using just the maze task time for completion and the error score, they were able to accurately classify 77.4% of the sample into Pass and Fail groups, with a sensitivity of 84.0% and a specificity of 61.8%. This is an impressive classification using just two scores from a single test and with a substantial sample size for this population. Unfortunately the author did not submit the model to statistical procedures to determine its stability and generalizability.

Lincoln et al. (2006) recruited 37 adults with dementia and a control group to compare the results of cognitive testing on on-road driving assessment that resulted in ratings of ‘probably safe’ as well as ‘probably unsafe’ and ‘definitely unsafe’. Fortunately, they constructed a classification model with just the dementia sample. They used discriminant analysis, which is very similar to BLR but is for use with parametric data. Discriminant analysis suffers from the same problems with overfitting as BLR and, unfortunately, Lincoln et al. offered 13 measures to their model for a variable to participant ratio of approximately 1 to 3 which likely led to substantial overfitting of their model. The authors reported an overall classification of 92.0% with a sensitivity of 90.0% and a specificity of 93.0%. Their study was the only one found to have estimated the predictive accuracy of the classification, by recruiting 17 new participants with dementia with which they tested their model equation (see Table 8-22). This model had an overall accuracy of 58.8% with a sensitivity of 40.0% and a specificity of 66.7%. This large reduction in accuracy may be explained by the likely large amount of overfitting present in their classification model due to the large number of variables entered. Lincoln et al. went on to choose a new cut-point for their model based on cut-points provided by ROC analysis which resulted in a model with 100% sensitivity. This is poor practice as in order to determine the potential accuracy of the classification model on a new sample the same cut-point used in the original model must be used when estimating predictive accuracy.

Duchek et al. (1998) recruited a sample of 44 people with CDR scores of 0.5 and 1.0 and used multiple regression to classify error score on an on-road driving assessment.

The authors did not provide the age or sex profiles of their participants. Measures accepted into the BLR backwards stepwise classification model included visual search and divided attention measures as well as the Boston Naming Test (a measure of the ability to correctly name objects which is often impaired in dementia). Multiple regression provides a continuous outcome score which provides only a measure of variance accounted by the model, which was 0.65. No estimate of predictive accuracy was performed. Finally, Dawson et al. (2009) recruited 40 participants with Alzheimer's dementia (83% male) who completed a number of cognitive measures and an on-road driving assessment to classify a driving error score using multiple regression. The measures of age, errors on the Benton Visual Retention Test (which assesses visual perception, memory, and visuoconstructive abilities) and Trail Making Test A accounted for 23% of the variance in the on-road driving error score.

The study by Snellgrove (2000) provides the best comparison to the models produced in the current study. Their accuracy of classification is impressive given the inclusion of only two measures from a single test. The classification models from the current study have likely benefited from the addition of extra tests but are also more likely to have been affected by overfitting, which is reflected in the drops in all three models following leave-one-out cross-validation. The Lincoln et al. (2006) model is likely seriously affected by overfitting as shown by the large reduction following testing on a new sample, and the authors actions of selecting a new cut-point in order to improve this outcome suggests they lack understanding of how to validate a classification model. The Duchek et al. (1998) and Dawson et al. (2009) studies are not as easily comparable to the models of the current study due to their use of multiple regression to classify a continuous on-road error score rather than a binary Pass/Fail outcome. In both cases, however, the reported variance accounted for is lower than for the Dementia and Driving extended assessment model.

Table 8-21. Classification models-comparison of current and other studies with drivers with predominantly with MCI or dementia

Reference	Statistical Model	Measures in Model	Accuracy	Sensitivity	Specificity	Variance Accounted for (R^2)	n	Mean Age of Sample
Dementia and Driving (n=32) standard model (the current study)	Binary logistic regression	Heart disease, Category Fluency, ADAS-Cog Total score	93.8%[1]	90.9%[1]	95.2%[1]	0.72	32	76.2
Dementia and Driving (n=32) alternate model (the current study)	Binary logistic regression	Heart disease, Category Fluency, ADAS-Cog Total score	93.8%[1]	90.9%[1]	95.2%[1]	0.72	32	76.2
Dementia and Driving (n=60) (the current study)	Binary logistic regression	Ballistic Movement peak velocity, Complex Attention invalid trials, Complex Attention reaction time STD, Random Tracking	83.3%[1]	66.7%[1]	92.3%[1]	0.47	60	77.9
Snellgrove (2000)	Binary logistic regression	Maze task completion time and error score	77.4%	84.0%	61.8%	N/S[2]	115	77.1
Lincoln et al. (2006)	Discriminant analysis	Dot cancellation measures (x 2), square matrices measures (x2), road sign recognition, MMSE, Stroop, Behavioral Assessment of the Dysexecutive Syndrome (x2 subtests), VOSP Incomplete Letters, an object recognition test, information processing tests (x 2)	92.0%	90.0%	93.0%	N/S[2]	37	71.0
Duchek et al. (1998)	Multiple regression	Visual search and divided attention measures, Boston Naming Task	N/S[2]	N/S[2]	N/S[2]	0.65	44	NS[2]
Dawson et al. (2009)	Multiple regression	Age, Benton Visual Retention Test errors, Trail Making Test A	N/S[2]	N/S[2]	N/S[2]	0.23	40	75.1

[1]Percentages are for the cut-point with the highest average of sensitivity and specificity. [2]N/S = not stated and unable to be computed from provided data.

Table 8-22. Comparison of the current study's estimated predictive accuracy compared to the estimation of predictive accuracy of other studies using samples of predominantly MCI and dementia samples

Reference	Method of estimating predictive accuracy	Measures in Model	Accuracy	Sensitivity	Specificity	n	Mean Age of Sample
Dementia and Driving (n=32) standard model (the current study)	Leave-one-out cross-validation of binary logistic regression model	6 measures offered with most iterations utilizing Heart disease, Category Fluency, and ADAS-Cog Total score	71.9%[1]	54.6%[1]	81.0%[1]	32	76.2
Dementia and Driving (n=32) alternate model (the current study)	Leave-one-out cross-validation of binary logistic regression model	6 measures offered with most iterations utilizing Heart disease, Category Fluency, and ADAS-Cog Total score	75.0%[1]	54.6%[1]	85.7%[1]	32	76.2
Dementia and Driving (n=60) (the current study)	Leave-one-out cross-validation of binary logistic regression model	12 measures offered with most iterations utilizing Ballistic Movement peak velocity, Complex Attention invalid trials, Complex Attention reaction time STD, and Random Tracking	68.3%[1]	47.6%[1]	79.5%[1]	60	77.9
Lincoln et al. (2006)	New sample of 17 people to test the model developed on 37 people	Dot cancellation measures (x 2), square matrices measures (x2), road sign recognition, MMSE, Stroop, Behavioral Assessment of the Dysexecutive Syndrome (x2 subtests), VOSP Incomplete Letters, an object recognition test, information processing tests (x2)	58.8%	40.0%	66.7%	17	75.8

[1]Percentages are given for the cut-point in the classification model with the highest average of sensitivity and specificity. [2]Value derived from the values given for sensitivity, specificity and total sample size.

8.6 Review of Study Hypotheses

Given the results of the study, it is possible to address the hypotheses outlined in Section 4.3.

8. A combination of standard cognitive tests and *SMCTests* measures will provide on-road Pass and Fail accuracy statistics higher than that achieved by previous studies.

This hypothesis is not supported as the extended assessment model did not accept a combination of *SMCTests* measures along with cognitive measures but rather accepted a medical measure and two standard cognitive test measures. However, the model that was created had higher total accuracy, sensitivity, and specificity than Snellgrove's (2000) model (see Table 8-21). Comparisons with the Lincoln et al. (2006) model are not useful due to that model's expected large amount of overfitting caused by offering too many measures to the model. In any case, the three models from the current study fare better on all accuracy measures than the Lincoln et al. model.

9. Drivers with MCI or Alzheimer's dementia who report experiencing more distress from medical conditions will be more likely to receive a fail score in the on-road driving assessment.

This hypothesis was not supported as in the extended assessment sample there was no difference in the amount of distress reported by drivers in the on-road Fail group. In fact, there was a trend for the Pass group to report more distress which is in the opposite from expected direction. The existence of heart disease, however, was linked to increased Fail rates. Presumably heart disease has physical and/or cognitive sequelae which affect the driving task that are not better explained by distress caused to an individual. It is possible that drivers with dementia are not as distressed by physical illness due to the decrements in memory and personality changes which are characteristic of dementia. Unlike the Healthy Driver Follow-up study, we are unable to follow these drivers with dementia prospectively to determine whether distress caused by medical conditions was related to subsequent crashes and traffic offences as found in the previous study.

10. An on-road error list will contain errors which can discriminate between drivers with dementia or mild cognitive impairment who Pass and Fail an on-road driving assessment.

This hypothesis was supported since eighteen errors were found that discriminated between Pass and Fail on-road groups and most of these errors were also rated by the occupational therapist as being important contributors to her global judgment of Fail for many drivers. The two most frequently rated errors by the occupational therapist were ‘Decreased awareness of other road users’ and ‘Decreased awareness of environment’ which are not so much as errors as subjective opinions likely based on the performance of errors and other behaviours that occurred during the assessment. These two measures had the largest effect size differences between groups (1.00 and 0.93 respectively). The results of the error list can provide useful information about the types of errors likely to indicate a Pass or Fail outcome in a way that could be accepted as useful by driving assessors.

8.7 Summary

Sixty driving assessment centre referrals with mild cognitive impairment or Alzheimer’s dementia completed a computerized battery of sensory-motor and cognitive tests and a formal blinded on-road driving assessment. A backwards stepwise binary logistic regression model selected measures of reaction time and movement speed of the upper limbs, visuomotor planning and coordination, and sustained attention. Following leave-one-out cross-validation, this model was estimated to correctly predict 68.3% of an independent group of drivers with mild cognitive impairment and Alzheimer’s dementia into on-road Pass and Fail groups.

A subsample of 32 participants completed additional standard cognitive tests and provided information on medical conditions. A binary logistic regression model in this subsample was formed which selected measures of verbal fluency, the presence of heart disease, and a fairly comprehensive cognitive screen. Following leave-one-out cross-validation, this model would be expected to correctly predict 75.0% of an independent group of drivers with mild cognitive impairment and Alzheimer’s dementia into on-road Pass and Fail groups. The three measures in this model could be performed in around 35-50 min in a primary health setting. This model is preferable to the *SMCTests* only model.

CHAPTER 9 -

Concluding Summary and Outlook

9.1 Achievement of Objectives

The objectives outlined in the introductory chapter (Section 1.3) are reproduced below.

1. Review the effects of normal ageing on driving performance.
2. Review how mild cognitive impairment and dementia affect driving performance.
3. Review how driving ability is currently assessed.
4. Assess the accuracy of measures currently used to assess driving ability.
5. Determine measures that predict on-road driving performance in a group of older cognitively-unimpaired drivers.
6. Follow cognitively-unimpaired older drivers to determine if drivers who fail an on-road driving assessment are more likely to experience adverse driving events than drivers who pass the assessment.
7. Determine measures the predict on-road driving performance in a group of drivers with mild cognitive impairment and dementia.

The first four objectives were achieved by reviews of the literature presented in Chapters 2 and 3. Objectives 5 to 7 were achieved through completion of the Healthy Older Drivers (Chapter 6), Healthy Driver Follow-up (Chapter 7), and Dementia and Driving (Chapter 8) studies.

9.2 Main Findings

The Healthy Older Drivers study produced a binary logistic regression model that included two measures (Trail Making Task B and Random Tracking run 1) that classified on-road Pass and Fail groups for cognitively-unimpaired drivers with 73.3% accuracy at the cut-point with the highest mean of sensitivity and specificity. However, the positive predictive values of 50.0% for the Optimized cut-point and 31.8% following leave-one-out cross-validation indicates that just over two thirds of people who would be predicted to Fail the on-road assessment would actually Pass. This high false positive rate would lead to many drivers being referred for unnecessary on-road assessments and the model cannot be recommended for use in this population.

The Healthy Driver Follow-up study found that there was no increase in negative driving outcomes for those who Failed the on-road assessment as part of the Healthy Older Drivers study. Albeit with a small sample size, this unique study suggests that on-road assessments are not useful in determining future real-world negative driving outcomes in drivers without diagnosed cognitive impairment. Our result is in agreement with a reassessment of the findings of Keall and Frith (2004a) and Anstey et al. (2009) who are the only other researchers to have followed a group of predominantly cognitively-unimpaired older drivers in order to determine the predictive ability of an on-road driving assessment outcome for future real-world negative driving events. The Healthy Driver Follow-up study also detected significant reductions in long distance driving as well as bicycle use, and in walking as a form of transportation over a two year period. This behaviour change measured longitudinally is not subject to cohort effects that are a limitation in cross-sectional designs.

The Dementia and Driving study comprised a sample of drivers known to be at increased risk for unsafe driving. Large effect size differences between on-road Pass and Fail groups for *SMCTests* measures, cognitive measures, and medical conditions provided ample scope to construct classification models. The standard *SMCTests*-only model did not have high enough accuracy statistics to recommend its use, with sensitivity scores across the two cut-points ranging from 48% to 52% and positive predictive values that would mean around half of drivers predicted to Fail the assessment would actually Pass. This could be due to the novelty of the *SMCTests* testing scenario which may be particularly difficult for people with MCI and Alzheimer’s dementia due to difficulties with learning new tasks.

The extended assessment model included cognitive and medical condition measures in preference to *SMCTests*. With resultant sensitivity of 55% for both cut-points following leave-one-out cross-validation, and positive and negative predictive values in the 60% to high 70% range, this model may offer a useful screen for determining whether drivers with MCI or Alzheimer’s dementia are more likely to Fail an on-road assessment. The inclusion of the ADAS-Cog in the extended assessment model supports recommendations (Iverson et al., 2010) that a rating of overall cognitive impairment (such as the CDR) can be useful in determining whether a driver with dementia is safe to continue driving. The extended assessment model could be used in primary health settings to assess the driving ability of patients with MCI and Alzheimer's dementia, and would take 40-50 minutes to perform.

Surprisingly, scores on the functional impairment questionnaires were not related to on-road driving outcome, with two of the questionnaires having effects in the opposite from expected direction which suggests they are not useful for guidance in deciding whether a driver with MCI or Alzheimer's dementia is safe to continue driving. Perhaps direct assessment of functional activities would be a better measure but such assessments require competent administration and scoring by occupational therapists and, hence, would make these tests unacceptable as a general screen for driving ability in this population.

The driving error list constructed for the Dementia and Driving study found many errors with high levels of agreement between observers and which also successfully discriminated between Pass and Fail groups. The Dementia and Driving study showed the inability of questions about road rules to discriminate between on-road Pass and Fail groups, which indicates an area that may not be worth assessing as part of formalized protocols for assessing driving ability in drivers with MCI and Alzheimer's dementia.

Variances in performance of the driving task in people with dementia has proven difficult in understanding driving performance. This is likely due to the nature of driving which is an over-learned task that may not compare well to the ways we seek to measure it using cognitive and sensory-motor measures. Given that a primary effect of dementia is to reduce a person's ability to learn new information and skills, it is perhaps not surprising that many participants perform very poorly on off-road assessments, or are unable to complete them, and yet can still drive to a standard considered safe by a driving assessor. The fact that the ADAS-Cog and Category Fluency proved useful in the extended assessment model in preference to the computerized *SMCTests* battery suggests that standard cognitive testing could be preferred over novel computerized or simulator tasks, even when the hardware used in those tasks are designed to be as face valid and familiar as possible.

The current studies are important for our understanding of both New Zealand drivers and older drivers in general. The results allowed for the writing of recommendations for the assessment of older drivers for both medical practitioners and occupational therapists as presented in Sections 9.4 and 9.5. The findings should also be of interest to those charged with implementing driving policy in New Zealand: the Ministry of Transport and the New Zealand Transport Agency.

9.3 Critique of the Project

There are always limitations to projects, especially those which attempt to deal with 'real-world' issues. Some of these limitations are examined below.

9.3.1 Sample Size

The construction of stable models that can be generalized to new samples is reliant on the recruitment of an adequate sample size. In order to be able to detect large effects sizes between Pass and Fail groups, with around one third of participants expected to fail the on-road assessment, we determined via power analysis that at least 60 participants would be required for each study. Sixty participants were successfully recruited for the Healthy Older Drivers study and we were able to find useful information regarding effect size differences on measures between Pass and Fail groups. In the Dementia and Driving Study, we initially aimed for all 60 participants to undergo the extended testing regimen which consisted of two additional appointments for collection of data additional to that provided by the standard *SMCTests* off-road test and the on-road driving assessment. It quickly became evident that many referrals to the driving assessment service were not interested in completing extra testing and the scope of recruitment was broadened to include a group of participants who were happy to have the results of their prescribed driving assessment included in a research study but who did not wish to undergo additional testing. It took 22 months of recruitment and testing to obtain data on 60 participants, with just over half agreeing to extended testing. Our power for finding large effects sizes was still adequate for the standard assessment sample but was lower for the extended assessment sample. Despite this, we were able to construct a model for the extended assessment sample whose accuracy was maintained relatively well following leave-one-out cross-validation. Even so, this model is unlikely to be as stable as the standard assessment model as it consists of fewer participants from whom to generalize. This was shown when two models were formed for the extended assessment sample, with leave-one-out cross-validation producing slightly different accuracy statistics based on a small difference in the six independent variables that were offered to construct the models. All participants for the Dementia and Driving sample were referrals to a driving assessment centre and are therefore likely to represent the kinds of people referred to other similar driving assessment services.

9.3.2 Estimated Predictive Accuracy of Classification Models

A discussion regarding the use of regression models is provided in Section 6.1.3.2. A critique of regression methods (Babyak, 2004) suggests that classification models should not be built on the basis of measures found to be related to the outcome variable in the sample. Rather, researchers are advised to recruit a second sample of participants to form models based on relationships found in a previous sample (or a previous study) and to enter a preset number of variables via forced entry. We did not construct our models in this way since we were interested in including measures that had not been properly investigated in these samples, namely *SMCTests*, certain cognitive measures, and personality measures. Neither did we enter variables at a ratio of at least 1 to 10 as recommended by Babyak, but instead chose a ratio of 1 to 5 as deemed acceptable by Tabachnick and Fidell (2001). We also used a step-wise elimination procedure to construct the model, something Babyak (2004) does not endorse. Two of the main criticisms of stepwise methods concern overfitting that occurs by offering too many variables to the model (Babyak, 2004) (a criticism that also applies for forced entry models when sample sizes are small), and the notion that the order in which variables are accepted into a forward stepwise model tells us something about the usefulness of those variables (Thompson, 1995) (i.e., the first measure accepted is the best predictor of the outcome variable). We minimized both of these concerns by a-priori selecting the number of variables that would be offered to each model at a ratio of 1 variable to 5 participants and by using a backwards elimination procedure with no attempt to determine which accepted variables were the ‘best’. In addition to these considerations, we also culled a number of variables that we determined did not have enough evidence to support their offering to the model or that we were only interested for descriptive purposes. This was done in order to concentrate on the variables we thought were mostly likely to be associated with driving in a predictable and explainable way. We also used collinearity diagnostics to detect variables that shared sufficient variance with one or more other measures in the model and could therefore be excluded from analysis. Most importantly, we used leave-one-out cross-validation to test the generalizability and stability of the model rather than relying on the results of classification which are likely overfitted to the sample.

The study by Innes et al. (in press) described in Section 6.3.4 found that while computationally advanced modelling techniques such as product kernel density and support vector machines produced classification models on on-road pass and fail outcomes that were almost 100% accurate, these models all fared around the same as more commonly used methods of discriminant analysis and binary logistic regression following leave-one-out cross-validation. These results highlight the importance of looking beyond classification to estimate the predictive accuracy for the models we construct. This is especially important when the results of our efforts could be used to determine important outcomes such as deciding on a person's ability to drive safely. We believe the methods used for this research project were better than most published driving research studies and we will continue to look at ways in which our techniques can be altered to provide a better description of driving behaviour.

9.3.3 Diagnosis of Dementia and MCI

Participants who took part in the extended assessment process of the Dementia and Driving study undertook substantial additional testing on cognitive and other measures. These measures were carefully selected both to provide evidence to assign diagnoses of MCI and Alzheimer's dementia, and also to explore the domains of cognition that were candidates for the prediction of driving ability. We did not have access to medical records or CT or MRI brain scans and were relying entirely on the cognitive testing, information from informants, and self-reported medical conditions in order to make MCI and Alzheimer's dementia diagnoses. It is possible that some of those in the Dementia and Driving study were incorrectly diagnosed. However, we are confident that our method of assessment was justified for the research project and it has provided a better understanding of the diagnostic status of older drivers referred for an on-road assessment due to concerns about declining cognitive skills, particularly since many referred drivers had not had any formal dementia assessment. Driving assessors must make their judgments of driving safety based on the often tentative diagnoses of "cognitive impairment" or "memory problems" that they receive on referral forms. Given this, it was reassuring to find that only a few recruited participants were subsequently excluded from analysis due to the decision that neither MCI or Alzheimer's dementia were present. This indicates that fairly accurate expectations of the presence of MCI and Alzheimer's dementia can be made on the basis of the referral letters, at least at the driving assessment centre that was used in the current studies.

9.3.4 Measures Not Included in the Current Studies

There may have been tests not used in the current studies (including ones currently available and ones not yet developed) that could have been useful for prediction of driving ability.

The Useful Field of View has been found to be useful in classifying self-reported crashes (De Raedt & Ponjaert-Kristoffersen, 2000; Hoffman & McDowd, 2010), officially-reported crashes (Owsley et al., 1991; Ball et al., 1993; Owsley et al., 1998; Sims et al., 1998; Ball et al., 2006), driving cessation (Ackerman et al., 2008; Edwards, Bart et al., 2009), and on-road driving assessment performance (De Raedt & Ponjaert-Kristoffersen, 2000; Whelihan et al., 2005; Stav et al., 2008; Uc et al., 2009), making it the most researched off-road computerized test. The most useful subtest of the UFOV is one measuring divided visual attention and appears to be useful even in drivers with no diagnosed cognitive impairment. This suggests it is sensitive to the effects of normal ageing on visual search, visual attention, and processing speed skills. The *SMCTests* measure Divided Attention is probably the closest comparison to the UFOV in the current project. In the Healthy Older Drivers study the measures in the Divided Attention test only had a small effect size difference between Pass and Fail groups, suggesting that the test is not sensitive enough to detect the effects of normal ageing. In the Dementia and Driving study the Divided Attention measure had moderate significant effects (for two of the three measures) which suggest it was sensitive enough to detect driving problems in this group. However, only one Divided Attention measure (number of arrows correct) was offered to the Dementia and Driving standard assessment model and was not accepted in the final model. Another Divided Attention measure (omission of arrows response) had a higher effect size (.76 versus .56 for number of arrows correct) but was deleted during collinearity tolerance testing due to sharing significant amounts of variance with one or more (unknown) variables. Duchek et al. (1998) found that many people with a CDR rating of 1 (a dementia severity rating of mild) found the UFOV too difficult to complete. Perhaps the *SMCTests* Divided Attention measure is easier than the UFOV as evidenced by most of the Dementia and Driving sample being able to complete the test, but the test not being a useful predictor in the Healthy Older Drivers group. Therefore, the UFOV may have been useful for inclusion in the Healthy Older Drivers study and its follow-up but maybe not as useful as the Divided Attention test in the Dementia and Driving study.

The Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005) is a brief screening tool similar to the MMSE which was specifically designed to be a more sensitive detector of the cognitive problems associated with MCI. The MoCA has been found to also be more sensitive to the detection of mild dementia than the MMSE (Smith et al., 2007). The MoCA could have been a useful substitute for the MMSE in the current studies and should be considered as a replacement for the MMSE in future driving research.

The Clock Drawing Test (CDT) is sensitive to the presence of Alzheimer’s dementia (Lezak et al., 2001). It consists of a person being asked to draw a clock face from memory with the hands of the clock pointing to a particular time, often 10 minutes past 11. In 2006 Molnar et al. (2006) performed a systematic review of the driving literature to find cognitive tests that were related to the driving ability of people with dementia. They found no studies that used the CDT, despite the test being recommended by the American Medical Association as a screening test for the detection of possible driving problems in people with dementia. Three studies by Freund and colleagues have investigated the usefulness of the CDT in samples of older drivers referred for driving assessments using simulated drive outcome measures. One study found a moderate correlation between CDT score and the number of errors performed (Freund et al., 2005). Another study found the CDT was the strongest classifier out of a group of tests for instances of unintended acceleration (Freund et al., 2008). Another study found significant differences in CDT score between groups rated as unsafe, safe, and restricted based on a simulated drive in which the assessors were not blinded to cognitive test outcomes (Freund & Colgrove, 2008). De Raedt and Ponjaert-Kristoffersen (2001) found that the CDT was accepted into a model that classified older drivers referred for a driving assessment into on-road pass and fail groups. Freund et al. (2008) consider the CDT to primarily be a measure of executive functioning, as well as measuring aspects of visuospatial and constructional abilities, memory, and abstract thinking. This claim for the CDT to be used as a measure of executive function is surprising, given that there are so many other more specific measures available which do not rely on intact visuospatial skills, such as the Trail Making Test, Colour-Word Interference, and Verbal Fluency measures. In any case, the CDT could be a useful addition to future driving research, particularly in investigating its use in detecting the on-road driving ability of those with dementia.

A measure of sensation-seeking may have been useful for the Healthy Older Drivers and Healthy Driver Follow-up studies as higher scores on sensation-seeking scales have been associated with higher rates of self-reported and simulated risky driving behaviour in college students (Schwebel et al., 2006) and higher numbers of driving violations and tickets in drivers aged 75 and over (Schwebel et al., 2007). Including a sensation-seeking scale was considered for the Healthy Older Drivers study but, as stated in Section 3.4, no scale specifically developed for older adults could be found. Since the items on current sensation-seeking scales may not necessarily translate to an older adult sample we were not confident that data from these scales would necessarily be valid. Personality variables were not included in the Dementia and Driving study due to the time needed to fill in forms when the extended assessment already required around five hours to conduct.

It would be useful to have a measure of the level of insight a person with cognitive impairment has into the effects of their condition on their driving behaviour, particularly since awareness of other road users and the environment were the two top rated reasons for drivers in the Dementia and Driving study failing the on-road assessment. A person with intact insight could be expected to monitor their driving performance and alter their behaviour to avoid complex driving scenarios situations. Unfortunately, insight is a difficult construct to measure, largely because it generally requires another person to rate the accuracy of a person's self-awareness which in itself is a highly subjective process. This is the system used by the DriveAware questionnaire (Kay et al., 2009a; Kay et al., 2009b; Kay et al., 2009c), but this questionnaire has not been used on its own to predict on-road outcome, only in combination with the DriveSafe test. Relying on the reports of significant others can be a problem due to denial of impairments, not wanting to hurt a person's feelings (especially if a significant other is asked to comment while the person in question is present), and deliberate minimization of difficulties in order to retain the status quo. Insight is affected by MCI and dementia but, as with the progress of all other cognitive impairments, it is difficult to predict and likely even harder to assess for the reasons outlined above. An attempt was made to measure insight in the Dementia and Driving study by having the occupational therapist rate the extent of insight for each participant (response of 'yes', 'limited', and 'no'). (see Appendix J for a copy of the form). No instruction was given to the occupational therapist regarding the meaning of 'insight'. As shown descriptively in Table 8-19 participants in the on-road Fail group were more often rated as having no insight in comparison to those in the Pass group. The statistics closely match the values given for whether participants had an awareness of the driving problems they performed on road, and in reality both of these measure could rely on each other. Also, both of these ratings could be affected by memory impairments, perhaps with participants who had forgotten the errors they performed more likely to be rated as having no insight or awareness of their driving problems as well as more likely to Fail the on-road assessment. Ratings of insight could have a place in driving assessment but the difficulty in defining what insight is and developing an appropriate way to measure it will be a barrier to its use.

The effect of age-related visual decline or medical eye conditions on driving was not a focus of the current study. Visual acuity in the left and right eye individually and together were included as measures in the Healthy Older Drivers study, but no significant associations were found between on-road Pass and Fail groups. Because of this, visual acuity was not pursued as a variable in the Dementia and Driving study, although all participants were screened for visual acuity at the off-road assessment and were found to meet mandatory New Zealand requirements for private motor vehicle licensure: minimum visual acuity of 6/12 using both eyes together, with or without correcting lenses. Both participant samples had a high percentage of people with cataracts, but medical conditions were assessed 12 months after the on-road assessment in the Healthy Older Drivers study and therefore could not be used in classifying their on-road Pass and Fail outcome. In the Dementia and Driving study eye conditions were not associated with on-road Pass and Fail outcome. In fact, more people in the Pass group endorsed having cataracts than in the Fail group (28.6% versus 9.1%). Anstey et al. (2005) found little evidence of the usefulness of visual measures to predict either on-road assessment outcome or real-world crash outcomes. Researchers may do better to concentrate on the more cognitively demanding aspects of vision such as measured by the UFOV. It is possible, however, that an interaction between cognitive and visual impairments could be more useful for predicting driving ability than visual or cognitive measures on their own.

The driving error list in the Dementia and Driving study had space for the assessors to record whether a participant was able to alter specific poor driving behaviours during the remainder of their driving assessment. This measure was not analyzed since there would not have been equal opportunities for all participants to perform a manoeuvre again during the remainder of their drive. Knowing whether drivers with MCI and dementia can learn to change their driving behaviour would be immensely helpful in determining their risk of unsafe driving. For example, only one driver in the Dementia and Driving study was rated as being able to improve their driving with lessons, and lessons were not recommended to this person. This likely indicates the occupational therapist does not believe that drivers with dementia can benefit from lessons. This conclusion would make intuitive sense given that a major deficit in dementia is in learning new information, but we do not know this for sure. Any on-road measure of whether driving errors were modified would not be complete without a study investigating whether any changes made persist beyond the driving assessment itself. There is no use in a driving behaviour changing during the assessment drive but then changing back during real-world driving, particularly if this is due to the instruction being forgotten. If it was found that some drivers with dementia retained the ability to modify their driving behaviour, then this could potentially be a useful measure for determining driver safety.

9.3.5 Reliability and Validity of On-Road Assessment

In the Healthy Older Drivers study, 27% of the participants failed the on-road assessment, compared with 35% of participants in the Dementia and Driving study. This seems a remarkably small difference between samples, especially since the Dementia and Driving sample were diagnosed with cognitive impairments necessary to meet requirements for MCI and Alzheimer’s dementia diagnoses. In terms of failure rates in other studies with dementia samples, these range from 27% to 63% (Fox et al., 1997; Grace et al., 2005; Lincoln et al., 2006), which is in keeping with the fail rate found in the Dementia and Driving sample.

In terms of failure rates in samples with healthy older drivers, Keall and Frith (2004a) found that 22% of a general population sample of over 39,000 drivers aged 80 and over failed a first attempt at an on-road driving test. In a sample of 270 predominantly cognitively-unimpaired drivers aged 70 and over, Wood et al. (2008) found that 17.4% received an on-road score that recorded the presence of critical driving errors. The Fail rate in the Healthy Older Drivers study is higher than both the Keall and Frith and Wood et al. studies. This could be due to a number of reasons. Firstly, inter-rater differences in driving assessors ratings between studies could influence Fail rates. Secondly, drivers in the Healthy Older Drivers study may not have driven as carefully during the study driving assessment as they would have had the outcome of the assessment been enforced, i.e., if a Fail rating led to licence cessation or the requirement for driver training. It is also possible that the lack of consequences for the on-road assessment in this study led to the occupational therapist giving more Fail ratings than she would have had the drivers faced consequences. However, drivers in the Wood et al. study also suffered no consequences from their on-road assessments which dilutes the influence of these two factors. An inspection of the brief reports provided for the 16 participants who failed in the Healthy Older Drivers study locates many errors which appear to support Fail scores, such as driving on the wrong side of the road, excessive speeding, failing to give way to cars or pedestrians, and inability to negotiate double-lane roundabouts. For many participants a Fail score appears to be justified, but it is possible that at least some were judged more harshly than they should have been. It may simply be the case that many cognitively-unimpaired older drivers make mistakes that would lead to on-road assessment Fail scores (something we could expect to see in drivers of all ages). The majority of drivers with MCI and mild Alzheimer's dementia may indeed be safe to continue driving. Of the 32 participants for whom dementia severity was estimated, only six were in the moderate range. Since the majority of the 32 people in the extended assessment were either MCI or mild Alzheimer's dementia it may be expected that only 35% received an on-road Fail rating.

As discussed in Sections 3.8 and 6.3.4, the on-road assessments used in both the Healthy Older Drivers and the Dementia and Driving studies had not been tested for reliability or validity. We attempted to address some of these issues in the current research project, including construction of an on-road error list from the errors performed during the Healthy Older Drivers study for detection of errors in the Dementia and Driving study. Additionally, in the Healthy Driver Follow-up we were able to show that on-road Pass/Fail outcome was not related to subsequent crashes and traffic offences. In the Dementia and Driving study, we measured the inter-rater reliability of detection of errors on the on-road assessment and found high reliability ratings for many errors. Since the classification models in this research project were trained and tested against the on-road Pass and Fail outcome, our models were critically dependent upon the accuracy, reliability, and validity of our 'gold standard' on-road assessment. We are by no means alone in this situation, and ways for addressing this are outlined in Section 9.6 below.

9.4 Recommendations for Medical Practitioners

9.4.1 Prioritizing Access to Driving Assessment Resources

Results from the Healthy Older Drivers and Healthy Driver Follow-up studies provide evidence to support a recommendation that on-road assessments are not necessary for older drivers who do not have diagnosed cognitive impairments. As discussed in Chapter 2, cognitively-unimpaired older drivers are amongst the safest drivers on the road and do not pose an increased risk to other road users. To counteract their increased physical fragility, older drivers can decrease their risk of being injured or killed on the road by purchasing cars with multiple safety features in order to reduce the change of injury should a crash occur.

9.4.2 Cognitive Impairment and Driving

The results of the literature review and the Dementia and Driving study have shown that people with Alzheimer's dementia who Fail an on-road assessment have significantly worse scores, at a group level, on a number of cognitive and sensory-motor tests than those who Pass (since only one participant with MCI failed the on-road assessment we cannot state the same finding with confidence for this group). It is therefore vital that people with cognitive impairments are detected and diagnosed in order that their risk can be assessed.

General practitioners are in the best position to screen for cognitive impairments in their older patients, and in New Zealand the compulsory licence renewal ages of 75, 80, and every two years thereafter are ideal times for cognitive status to be assessed. We do not recommend the Mini-Mental State Exam for screening since it will miss many people with MCI and mild dementia. Instead, the Montreal Cognitive Assessment (MoCA) is a similar short screen with higher sensitivity for detecting MCI and mild dementia than the Mini-Mental State Exam (Nasreddine et al., 2005; Smith et al., 2007). The ADAS-Cog proved useful in the Dementia and Driving study, although it can take around 40 minutes to complete for a person with dementia and may need to be performed by a practice nurse rather than a doctor because of this. The detection of possible cognitive impairment using these measures is only the first step of a diagnosis and a diagnosis of dementia does not mean that a person is an unsafe driver. Everyone with dementia will have to stop driving at some point, with several researchers suggesting assessment every six months (Fox et al., 1997; Dubinsky et al., 2000; Duchek et al., 2003; Adler et al., 2005; Frittelli et al., 2009). Physicians are recommended to follow guidelines that that drivers with moderate or severe dementia cease driving (Canadian Medical Association, 2006; Australian and New Zealand Society for Geriatric Medicine, 2010).

9.4.2.1 Assessment of Driving History

Physicians are encouraged to assess their patients with dementia more frequently than the two-yearly compulsory licence renewal medical evaluations that occur in New Zealand from age 80. It is worth taking more than a cursory approach to this questioning and there are a few guidelines to help with this. The reported absence of recent crashes and traffic offences does not necessarily mean that a person is a safe driver. Crashes are low base-rate occurrences, even for impaired drivers, and other road users are often able to avoid unsafe drivers. Inquiring about ‘discussions’ with police officers, whether or not an infringement notice was administered, or whether other drivers have been complaining about the driver’s behaviour may provide useful information.

A caregiver’s rating of a patient’s driving ability as marginal or unsafe has been shown to be related to adverse on-road outcomes but a patient’s self-rating of driving ability is not (Iverson et al., 2010). Some family members may be reluctant to talk about the issue of driving, particularly in the presence of the affected family member. Family members may also be invested in allowing an older relative to drive despite increased risk. It is important to remember that holding a driver licence is a privilege and not a right and allowing an unsafe driver to continuing driving puts the patient, their passengers, and other road users at increased risk of being injured or killed on the road.

9.4.2.2 Driving Cessation

When a person must immediately cease driving, common sense actions are required. Family or supporters may need to take responsibility for access to car keys, or even disabling or removing a vehicle in some situations. When assessment of driving is delayed, driving may need to cease in the interim. Informal management of driving cessation is preferred but the threshold for deciding when to repeat assessment must be low since the progress of cognitive deterioration can be quick. Decline in cognitive function and reports of increased problems with driving from significant others should prompt further assessment of driving ability.

9.4.2.3 Suggested Process for Determining Driver Safety for Older Drivers

A flowchart has been constructed that proposes a pathway for medical practitioners to navigate when making decisions on (i) older patients presenting for licence renewal or (ii) patients for whom possible or definite cognitive impairment has been raised by the patient or a family member or observed by the practitioner (Figure 9-1). The flowchart addresses only the impact of MCI and Alzheimer’s dementia. Medical practitioners must also consider the affects of medical conditions including other neurological disorders, psychiatric disorders, musculoskeletal disorders, and medications when making decisions regarding licence renewal.

The flowchart makes useful of the Montreal Cognitive Assessment (MoCA) for initial screening of cognitive processes, with further more detailed examinations to be undertaken to make a diagnosis of MCI or Alzheimer's dementia. It is recommended that drivers with moderate or severe dementia cease driving. For those with MCI or mild dementia there are two options to choose from. If medical driving assessments are available in the region where the practitioner is based then the patient should be referred for an assessment (in the context of New Zealand, this should be a medical driving assessment and not the On-Road Safety Test).

If medical driving assessments are not available, or if the referral is rejected, or if the patient cannot afford to pay for a private assessment, then the model developed in the Dementia and Driving study to determine the patient's risk of failing an on-road assessment should be used. Medical practitioners or their practice nurses would need to be adequately trained in administration and scoring of these tests by an appropriately qualified person (i.e., psychologist or psychiatrist). The binary logistic regression equation can be performed by hand following collection of the scores for the three measures (heart disease is rated as a binary 1 if present and 0 if absent). The cut-point with the optimal balance of sensitivity and specificity in the classification model should be used. If a patient scores 0.53 or higher then they are predicted to Fail the on-road assessment. If the score is less than 0.53 then they are predicted to Pass. Using this cut-point, a Fail rating should be accurate 60% of the time (40% of the time the patient would be expected to Pass an on-road assessment), and a Pass rating will be accurate 77.3% of the time (23% of the time the patient would be expected to Fail the on-road assessment).

If the patient is predicted to Pass, then the practitioner can renew their licence but must monitor the patient at regular intervals for signs of cognitive deterioration. The two-yearly compulsory renewals for drivers aged 80 and over if not frequent enough to assess the impact of dementia on driving. If the patient receives a predicted Fail score then the practitioner is advised to recommend driving cessation to their patient and to inform the New Zealand Transport Agency of this decision.

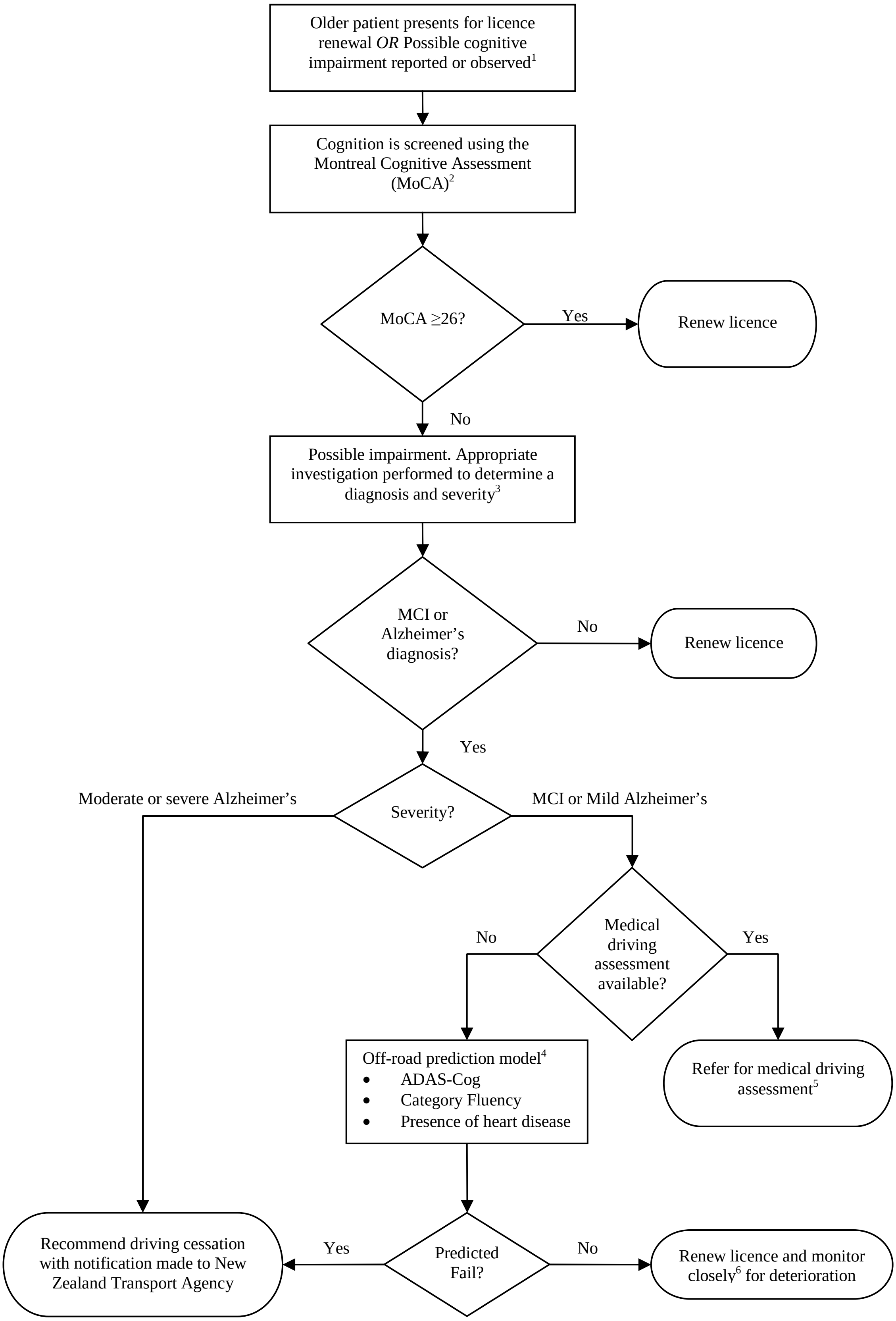


Figure 9-1. Decision pathway for medical practitioners for older patients presenting for licence renewal and for those with possible or of definite Alzheimer's dementia or MCI

[1]The flowchart addresses only the impact of MCI and Alzheimer’s dementia. Medical practitioners must also consider the affects of medical conditions including other neurological disorders and medications when making decisions regarding licence renewal.

[2]Test forms and administration instructions for the MoCA are available from the MoCA website http://www.mocatest.org/

[3]Diagnosis of Alzheimer’s dementia can be made using DSM-IV-TR criteria (American Psychiatric Association, 2000) or NINCDS-ADRDA criteria (McKhann et al., 1984). Both criteria sets require the presence of impairment in two cognitive domains which must include memory, decrements in activities of daily living, and no other primary reason for the cognitive impairment. MCI criteria are available from Petersen (2004) and must include one or more cognitive impairments, which do not have to include memory, and no significant impairment in activities of daily living.

[4]The binary logistic regression equation can be performed by hand following collection of the scores for the three measures (heart disease is rated as a binary 1 if present and 0 if absent). The cut-point with the optimal balance of sensitivity and specificity in the classification model should be used. If a patient scores 0.53 or higher then they are predicted to Fail the on-road assessment. If the score is less than 0.53 then they are predicted to Pass. Using this cut-point, a Fail rating should be accurate 60% of the time (40% of the time the patient would be expected to Pass an on-road assessment), and a Pass rating will be accurate 77.3% of the time (23% of the time the patient would be expected to Fail the on-road assessment.

[5]The On-Road Safety Test is not appropriate for a person with MCI or Alzheimer’s dementia.

[6]Some researchers suggest that on-road assessment for people with dementia be performed every six months (Fox et al., 1997; Dubinsky et al., 2000; Duchek et al., 2003; Adler et al., 2005; Frittelli et al., 2009). The compulsory licence renewal intervals of 2 years once a person is 80 years of age are too infrequent to assess the possible impact of Alzheimer’s or MCI on driving. Medical practitioners should actively seek information about driving at least every six months. Information from family members is a better measure of driving ability than information from the patient themselves (Iverson et al., 2010).

9.5 Recommendations for Driving Specialist Occupational Therapists

Based on the findings of the Dementia and Driving study, a number of recommendations are made for occupational therapists in New Zealand. Firstly, if on-road assessments are available it is recommended that assessors perform only an on-road assessment with no formalized off-road testing component. This will reduce the unnecessary stress of an off-road testing appointment. On-road assessment can start either from the testing centre or the patient’s home.

If on-road assessment is not possible then off-road assessment will need to be performed. *SMCTests* should not be used as an off-road assessment for drivers with MCI and Alzheimer’s dementia. Instead the model using the ADAS-Cog Total score, Category Fluency tests and the binary instance of heart disease should be used. Assessors will need to be adequately trained in administration and scoring of these tests by an appropriately qualified person (i.e., a psychologist or psychiatrist). Driving assessors should follow the same instructions for scoring as detailed for medical practitioners in Section 9.4.2.3.

Assessors may wish to reject referrals for people with moderate and severe dementia and instead instruct the referrer to recommend driving cessation. If a referral does not include a severity rating then this should be requested from the referrer.

Assessors should keep in mind that several research groups recommend 6-monthly assessments for drivers with dementia.

9.6 Future Directions

In terms of the field of driving research as a whole, the Healthy Older Drivers and Dementia and Driving studies have reinforced the importance of estimating predictive accuracy of classification models to determine the potential generalizability of the models as well as to provide better estimates of accuracy statistics which is not possible when models are not tested beyond classification. It is also important to recognise the problems associated with overfitting classification models, as estimating their predictive accuracy will likely result in large drops in accuracy statistics which belie the potential usefulness of the model had appropriate numbers of variables been offered for the sample size. Babyak (2004) is a useful resource for determining methods for reducing the impact of overfitting in regression models.

In addition to the leave-one-out cross-validation we completed, the classification models developed in the Dementia and Driving study should ideally be tested on a new sample of participants to confirm their generalizability. In a driving assessment setting, this would consist of performing the tests or collecting the health information data during the off-road portion of the driving assessment and then conducting an on-road driving assessment with the assessors blinded to the results of off-road testing. The off-road testing could be performed in around 40-50 min if using the extended assessment model.

A prospective follow-up of drivers with MCI and Alzheimer’s dementia who Pass an on-road assessment could be performed in order to measure the incidence of crashes and traffic offences, and the time until licence cessation. If this follow-up spanned several years, data relating to the outcomes of repeated driving assessments could provide information about changes in driving behaviour over time.

If the measures from the extended assessment Dementia and Driving models were adopted by general practitioners, a survey of the ease of administration and perceived usefulness of the screen would be worthwhile. This information could be used to assess the suitability and usability of the predictive model and provide direction for its implementation in additional primary health care services.

Measures not included in the current studies (as discussed in Section 9.3.4) would be useful for addition in future studies. These include the UFOV, the MoCA, measures of insight (perhaps DriveAware), and sensation seeking (provided appropriate measures could be found or developed), and the driving error list developed following the Healthy Older Driver study.

The most pressing issue for driving research is in the choice of outcome measures that indicate safe or unsafe driving. A variety of outcome measures are currently used: on-road driving assessment outcome (either pass/fail, pass/fail/conditional pass, or number of errors performed during the assessment), driving assessment outcome using a driving simulator (same outcome measures as for an on-road assessment), self-reported crashes or offences (retrospective or prospective), significant other-reported crashes or offences (retrospective or prospective) officially recorded crashes or offences (retrospective or prospective), and officially recorded at-fault only crashes (retrospective or prospective).

The validity of on-road driving assessments for predicting future adverse driving events is a major issue for driving researchers. On-road assessments are generally accepted as the ‘gold standard’ measure of driving ability and yet most assessments do not have standardized scoring systems or a standardized cut-point for determining Pass and Fail outcome (Korner-Bitensky et al., 2006). As long as the ability of on-road assessments to predict real-world driving behaviour is unknown, researchers will continue to build classification models that may not generalize outside of their study sample. There is a need for assessments to be tested for reliability (intra-rater, inter-rater, and test-retest reliability) and following this, tested for predictive validity for detecting those at risk of future adverse driving events, whether the outcome measure is Pass or Fail or the number of errors performed or types of errors performed during the assessment. This is an endeavour that must be developed in conjunction with researchers and the occupational therapists who will ultimately perform the assessments. There should also be standardized training programmes for occupational therapists given the task of making decisions about driving safety as decisions must be consistent and fair.

The ability of simulated driving measures to relate to real-world driving behaviour is for the most part unknown since most studies do not attempt to validate simulated driving outcomes against real-world driving outcomes. This, coupled with the expense of simulated driving systems, makes the widespread use of simulators in driving assessors centres unlikely, at least at present, and potentially limits the application of research findings into real-world service scenarios.

Self- and significant-other reported crashes and offences are subject to forgetting, lying, bias, or even malicious motives. Officially-reported crashes and offences are rare in comparison to self-report as they only sample a small amount of behaviour that can only hint at a person possibly being a less safe driver. Official crash data is biased to more serious crashes and, at least in New Zealand, to crashes where injury or death were involved only. The low base-rate of these kinds of crashes require for large samples of participants to be recruited and, as shown in the Healthy Driver Follow-up study, the differences in crash rates between drivers judged to have passed or failed an on-road assessment are not necessarily very different. By limiting crash data to at-fault crashes only further reduces the base-rate of crashes that can included in studies.

Underlying all measures that seek to define on-road safety is the necessity for these measures to be related to the real-world negative outcomes that we are seeking to reduce. An outcome measure that does not predict prospective real-world negative driving outcomes is not useful. The difference in real-world on-road outcomes between drivers rated as ‘safe’ and ‘unsafe’ also has to be substantial in order to avoid unnecessary licence cessation and the negative outcomes that proceed from this. If effect sizes between groups are small, we can argue that the clinical significance of findings is negligible, no matter what the *p* value. Establishing truly useful outcome measures for driving assessment is not an easy task, but may be the most important consideration for the field of driving research.

CHAPTER 10 -

References

Ackerman, M. L., Edwards, J. D., Ross, L. A., Ball, K. K., & Lunsman, M. (2008). Examination of cognitive and instrumental functional performance as indicators for driving cessation risk across 3 years. *The Gerontologist, 48*, 802-810.

Adler, G., Rottunda, S., & Dysken, M. (2005). The older driver with dementia: an updated literature review. *Journal of Safety Research, 36*, 399-407.

Aging Clinical Research Center (2011). Geriatric Depression Scale. Retrieved 6 February 2011, from http://www.stanford.edu/~yesavage/GDS.html

Alvarez, F. J., & Fierro, I. (2008). Older drivers, medical condition, medical impairment and crash risk. *Accident Analysis and Prevention, 40*, 55-60.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Anstey, K. J., Wood, J., Caldwell, H., Kerr, G., & Lord, S. R. (2009). Comparison of self-reported crashes, state crash records and an on-road driving assessment in a population-based sample of drivers aged 69-95 years. *Traffic Injury Prevention, 10*, 84-90.

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review, 25*, 45-65.

Australian and New Zealand Society for Geriatric Medicine (2010). Position statement: dementia and driving. *Australasian Journal on Ageing, 29*, 137-141.

Babyak, M. A. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine, 66*, 411-421.

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: expanding the useful field of view. *Journal of the Optical Society of America, A, Optics, Image & Science, 5*, 2210-2219.

Ball, K. K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology & Visual Science, 34*, 3110-3123.

Ball, K. K., Roenker, D. L., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin, G., Raleigh, R., Joyce, J. J., Cissell, G. M., & Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a Department of Motor Vehicles setting? *Journal of the American Geriatrics Society, 54*, 77-84.

Beck, A. T. (1990). *Beck Anxiety Inventory*. San Antonio: The Psychological Corporation.

Benton, A. L. (1983). *Judgment of Line Orientation*. Lutz, Florida: Psychological Assessment Resources, Inc.

Benton, A. L., Sivan, A. B., Hamsher, K. D., Varney, N. R., & Spreen, O. (1994). *Contributions to neuropsychological assessment: a clinical manual* (2nd ed.). New York: Oxford University Press.

Bowens, A. (2004). *Computerised off-road driving assessment*. Wellington: Accident Compensation and Rehabilitation Insurance Corporation.

Brainmetric (2011). Trailmaking Test part A and B. Retrieved 6 February 2011, from http://www.brainmetric.com/resource.htm

Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., Logan Jr., W. C., Ogle, J. H., Tyrrell, R. A., & Wills, R. F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis and Prevention, 42*, 788–796.

Brown, L. B., & Ott, B. R. (2004). Driving and dementia: a review of the literature. *Journal of Geriatric Psychiatry and Neurology, 17*, 232-240.

Brown, L. B., Stern, R. A., Cahn-Weiner, D. A., Rogers, B., Messer, M. A., Lannon, M. C., Maxwell, C., Souza, T., White, T., & Ott, B. R. (2005). Driving scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. *Archives of Clinical Neuropsychology, 20*, 209-215.

Canadian Medical Association (2006). *Determining medical fitness to operate motor vehicles: CMA driver's guide - 7th edition*. Ottawa, ON.

Christchurch Neurotechnology Research Programme (2006). *Canterbury Driving Assessment Tool (CanDAT™) incorporating SMCTests™ Version 5.0: User's Manual*. Christchurch, New Zealand: Christchurch Neurotechnology Research Programme.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Cooper, P. J. (1997). The relationship between speeding behaviour (as measure by violation convictions) and crash involvement. *Journal of Safety Research, 28*, 83-95.

Cooper, P. J., Tallman, K., Tuokko, H., & Beattie, B. L. (1993). Vehicle crash involvement and cognitive deficit in older drivers. *Journal of Safety Research, 24*, 9-17.

Dahlen, E. R., & White, R. P. (2006). The Big Five factors, sensation seeking, and driving anger in the prediction of unsafe driving. *Personality and Individual Differences, 41*, 903-915.

Davey, J., Wishart, J. F., & Watson, B. (2007). An application of the driver behaviour questionnaire in an Australian organisational fleet setting. *Transportation Research Part F, 10,* 11-21.

Dawson, J. D., Anderson, S. W., Uc, E. Y., Dastrup, E., & Rizzo, M. (2009). Predictors of driving safety in early Alzheimer disease. *Neurology, 72,* 521-527.

De Raedt, R., & Ponjaert-Kristoffersen, I. (2000). The relationship between cognitive/neuropsychological factors and car driving performance in older adults. *Journal of the American Geriatrics Society, 48,* 1664-1668.

De Raedt, R., & Ponjaert-Kristoffersen, I. (2001). Short cognitive/neuropsychological test battery for first-tier fitness-to-drive assessment of older drivers. *The Clinical Neuropsychologist, 15,* 329-336.

De Ronchi, D., Berardi, D., Menchetti, M., Ferrari, G., Serretti, A., Dalmonte, E., & Fratiglioni, L. (2005). Occurrence of cognitive impairment and dementia after the age of 60: a population-based study from Northern Italy. *Dementia and Geriatric Cognitive Disorders, 19,* 97-105.

Deffenbacher, J. L., Lynch, R. S., Filetti, L. B., Dahlen, E. R., & Oetting, E. R. (2003). Anger, aggression, risky behavior, and crash-related outcomes in three groups of drivers. *Behaviour Research and Therapy, 41,* 333-349.

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Swaim, R. C. (2002). The Driving Anger Expression Inventory: a measure of how people express their anger on the road. *Behaviour Research and Therapy, 40,* 717-737.

Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving anger scale. *Psychological Reports, 74,* 83-91.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System*. San Antonio: Harcourt Assessment, Inc.

Dobbs, A. R. (1997). Evaluating the driving competence of dementia patients. *Alzheimer Disease and Associated Disorders, 11, suppl 1,* 8-12.

Dobbs, A. R. (2005). The development of a scientifically based driving assessment and standardization procedures for evaluating medically at-risk drivers. Paper presented at the *Canadian Multidisciplinary Road Safety Conference XV*, Fredericton, NB, Canada.

Dobbs, A. R., Caprio Triscott, J. A., & McCraken, P. N. (2004). Considerations for assessment of medical competence to drive in older patients. *Geriatrics and Aging, 7,* 42-46.

Dobbs, A. R., Heller, R. B., & Schopflocher, D. (1998). A comparative approach to identify unsafe older drivers. *Accident Analysis and Prevention, 30,* 363-370.

DriveABLE™ (2011). Retrieved 29 January 2010, from http://www.driveable.com/

Dubinsky, R. M., Stein, A. C., & Lyons, K. (2000). Practice parameter: risk of driving and Alzheimer's disease (an evidence-based review). *Neurology, 54*, 2005-2211.

Duchek, J. M., Carr, D. B., Hunt, L., Roe, C. M., Xiong, C., Shah, K., & Morris, J. C. (2003). Longitudinal driving performance in early-stage dementia of the Alzheimer type. *Journal of the American Geriatrics Society, 51*, 1342-1347.

Duchek, J. M., Hunt, L., Ball, K. K., Buckles, V., & Morris, J. C. (1998). Attention and driving performance in Alzheimer's disease. *The Journals of Gerontology: Psychological Sciences, 53B*, P130-P141.

Edwards, J. D., Bart, E., O'Connor, M. L., & Cissell, G. (2009). Ten years down the road: predictors of driving cessation. *The Gerontologist, 50*, 393-399.

Edwards, J. D., Lunsman, M., Perkins, M., Rebok, G. W., & Roth, D. L. (2009). Driving cessation and health trajectories in older adults. *Journal of Gerontology: Medical Sciences, 64A*, 1290-1295.

Edwards, J. D., Perkins, M., Ross, L. A., & Reynolds, S. L. (2009). Driving status and three-year mortality among community-dwelling older adults. *Journal of Gerontology: Medical Sciences*, *64A*, 300-305.

Evans, L. (2000). Risks older drivers face themselves and threats they pose to other road users. *International Journal of Epidemiology*, *29*, 315-322.

Fitten, L. J., Perryman, K. M., Wilkinson, C. J., Little, R. J., Burns, M. M., Pachana, N., Mervis, R., Malmgren, R., Siembieda, D. W., & Ganzell, S. (1995). Alzheimer and vascular dementias and driving: a prospective road and laboratory study. *Journal of the American Medical Association, 273*, 1360-1365.

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

Fox, G. K., Bowden, S. C., Bashford, G. M., & Smith, D. S. (1997). Alzheimer's disease and driving: prediction and assessment or driving performance. *Journal of the American Geriatrics Society, 45*, 949-953.

Freeman, E. E., Gange, S. J., Munoz, B., & West, S. K. (2006). Driving status and risk of entry into long-term care in older adults. *American Journal of Public Health, 96*, 1254-1259.

Freund, B., Colgrove, L. A., Petrakos, D., & McLeod, R. (2008). In my car the brake is on the right: pedal errors among older drivers. *Accident Analysis and Prevention, 40*, 403-409.

Freund, B., & Colgrove, L. A. A. (2008). Error specific restrictions for older drivers: promoting continued independence and public safety. *Accident Analysis and Prevention, 40,* 97-103.

Freund, B., Gravenstein, S., Ferris, R., Burke, B. L., & Shaheen, E. (2005). Drawing clocks and driving cars. *Journal of General Internal Medicine, 20*, 240-244.

Freund, B., Gravenstein, S., Ferris, R., & Shaheen, E. (2002). Evaluating driving performance of cognitively impaired and healthy older adults: a pilot study comparing on-road testing and driving simulation. *Journal of the American Geriatrics Society, 50*, 1309-1315.

Frittelli, C., Borghetti, D., Iudice, G., Bonanni, E., Maestri, M., Tognoni, G., Pasquali, L., & Iudice, A. (2009). Effects of Alzheimer's disease and mild cognitive impairment on driving ability: a controlled clinical study by simulated driving test. *International Journal of Geriatric Psychiatry, 24,* 232-238.

Galasko, D., Bennett, D., Sano, M., Ernesto, C., Thomas, R., Grundman, M., & Ferris, S. (1997). An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. *Alzheimer Disease and Associated Disorders, 11*, S33-S39.

Garson, G. D. (2010). *Multiple Regression*. Retrieved 6 January 2010, from http://faculty.chass.ncsu.edu/garson/PA765/regress.htm

Gianutsos, R. (1994). Driving advisement with the Elemental Driving Simulator (EDS): when less suffices. *Behavior Research Methods, Instruments, & Computers, 26*, 183-186.

Gianutsos, R., Campbell, A., Beattie, A., & Mandriota, F. (1992). The Driving Advisement System: a computer-augmented quasi-simulation of the cognitive prerequisites for resumption of driving after brain injury. *Assistive Technology, 4,* 70-86.

Glantz, S. A. (2005). *Primer of Biostatistics* (6th ed.). New York: McGraw-Hill.

Grace, J., Amik, M. M., D'Abreu, A., Festa, E. K., Heindel, W. C., & Ott, B. R. (2005). Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. *Journal of the International Neuropsychological Society, 11*, 766-775.

Graham, D. P., Caully, J. A., Snow, A. L., Mass, P., & Doody, R. (2004). The Alzheimer's disease assessment scale - cognitive subscale. Normative data for older adult controls. *Alzheimer Disease and Associated Disorders, 18*, 236-240.

Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., Foster, N. L., Jack, C. R., Galasko, D. R., Doody, R., Kaye, J., Sano, M., Mohs, R., Gauhtier, S., Kim, H. T., Jin, S., Schultz, A. N., Schafer, K., Mulnard, R., van Dyck, C. H., Mintzer, J., Zamrini, E. Y., Cahn-Weiner, D., & Thal, L. J. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Archives of Neurology, 61*, 59-66.

Hagge, R. A. (1994).The California driver performance evaluation project: an evaluation of a new driver licensing road test. Sacramento, CA: California Department of Motor Vehicles. *Report No. 150,*

Hakamies-Blomqvist, L., Raitanen, T., & O'Neill, D. (2002). Driver ageing does not cause higher accident rates per km. *Transportation Research Part F, 5*, 271-274.

Herrera-Guzmán, I., Peña-Casanova, J., Lara, J. P., Gudayol-Ferré, E., & Böhm, P. (2004). Influence of age, sex, and education on the Visual Object and Space Perception Battery (VOSP) in a healthy normal elderly population. *The Clinical Neuropsychologist, 18,* 385-394.

Hoffman, L., & McDowd, J. M. (2010). Simulator driving performance predicts accident reports five years later. *Psychology and Aging, 25*, 741-745.

Hopkins, W. G. (2004). *A new view of statistics*. Internet Society for Sport Science. Retrieved 16 August 2004, from www.sportsci.org/resource/stats/

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley-Interscience.

Hunt, L., Morris, J. C., Edwards, D., & Wilson, B. S. (1993). Driving performance in persons with mild senile dementia of the Alzheimer type. *Journal of the American Geriatrics Society, 41*, 747-753.

Hunt, L. A., Murphy, C. F., Carr, D., Duchek, J. M., Buckles, V., & Morris, J. C. (1997). Reliability of the Washington University Road Test. *Archives of Neurology, 54*, 707-712.

Innes, C. R. H., Jones, R., Dalrymple-Alford, J., Severinsen, J., & Gray, J. (2009). Prediction of driving ability in people with dementia- and non-dementia- related brain disorders. *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 342-348.

Innes, C. R. H., Jones, R. D., Anderson, T. J., Hollobon, S. G., & Dalrymple-Alford, J. C. (2009). Performance in normal subjects on a novel battery of driving-related sensory-motor and cognitive tests. *Behavior Research Methods, 41*, 284-294.

Innes, C. R. H., Jones, R. D., Dalrymple-Alford, J. C., Hayes, S., Hollobon, S., Severinsen, J., Smith, G., Nicholls, A., & Anderson, T. J. (2007). Sensory-motor and cognitive tests can predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences, 260*, 188-198.

Innes, C. R. H., Lee, D., Chen, C., Ponder-Sutton, A. M., Melzer, T. R., & Jones, R. D. (in press). Do complex models increase prediction of complex behaviours? Predicting driving ability in people with brain disorders. *Quarterly Journal of Experimental Psychology*.

Iverson, D. J., Gronseth, G. S., Reger, M. A., Classen, S., Dubinsky, R. M., & Rizzo, M. (2010). Practice parameter update: evaluations and management of driving risk in dementia. *Neurology, 74*, 1316-1324.

Janke, M. K. (1991). Accidents, mileage, and the exaggeration of risk. *Accident Analysis and Prevention, 23*, 183-188.

Janke, M. K. (1994).Age-related disabilities that may impair driving and their assessment. Washington, DC: California Department of Motor Vehicles.

Janke, M. K., & Eberhard, J. W. (1998). Assessing medically impaired older drivers in a licensing agency setting. *Accident Analysis and Prevention, 30*, 347-361.

Johansson, K., Bronge, L., Lungberg, C., Persson, A., Seiderman, M., & Vittanen, M. (1996). Can a physician recognize an older driver with increased crash risk potential? *Journal of the American Geriatrics Society, 44*, 1198-1204.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: history, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: theory and research* (2nd ed., 102-138). New York: Guilford.

Jones, R. D. (2006). Measurement of sensory-motor control performance capacities: tracking tasks. In J. D. Bronzino (Ed.), *Biomedical Engineering Fundamentals* (3rd ed., Vol. 1, chap. 77: 71-25). Boca Raton, Florida: CRC Press.

Jones, R. D., & Donaldson, I. M. (1986). Measurement of sensory-motor integrated function in neurological disorders: three computerised tracking tasks. *Medical and Biological Engineering and Computing, 24*(536-540).

Jones, R. D., & Donaldson, I. M. (1995). Fractionation of visuoperceptual dysfunction in Parkinson's disease. *Journal of the Neurological Sciences, 131*, 43-50.

Jorm, A. F. (2004). The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): a review. *International Psychogeriatrics, 16*, 275-293.

Jorm, A. F., & Jacomb, P. A. (1989). The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine, 19*, 1015-1022.

Jurica, P. J., Leitten, C. L., & Mattis, S. (2001). *Dementia Rating Scale - 2*. Lutz, Florida: Psychological Assessment Resources, Inc.

Kay, L. G., Bundy, A., & Clemson, L. (2009a). Validity, reliabitility and predictve accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation, 31*, 1074-1082.

Kay, L. G., Bundy, A. C., & Clemson, L. (2009b). Awareness of driving ability in senior drivers with neurological conditions. *The American Journal of Occupational Therapy, 63*(146-150).

Kay, L. G., Bundy, A. C., & Clemson, L. M. (2009c). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation, 90*, 1514-1522.

Keall, M. D., & Frith, W. J. (2004a). Association between older driver characteristics, on-road driving test performance, and crash liability. *Traffic Injury Prevention, 5*, 112-116.

Keall, M. D., & Frith, W. J. (2004b). Older driver crash rates in relation to type and quantity of travel. *Traffic Injury Prevention, 5*, 26-36.

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., & Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *American Journal of Occupational Therapy, 60,* 428-434.

Korner-Bitensky, N., & Sofer, S. (2009). The DriveABLE Competence Screen as a predictor of on-road driving in a clinical sample. *Australian Occupational Therapy Journal, 56*, 200-205.

Lafont, S., Laumon, B., Helmer, C., Dartigues, J.-F., & Fabrigoule, C. (2008). Driving cessation and self-reported car crashes in older drivers: the impact of cognitive impairment and dementia in a population-based study. *Journal of Geriatric Psychiatry and Neurology, 21*, 171-182.

Langford, J., & Koppel, S. (2011). Licence restrictions as an under-used strategy in managing older driver safety. *Accident Analysis and Prevention, 43*, 487-493.

Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk - a replication of low mileage bias. *Accident Analysis and Prevention, 38*, 574-578.

Lau, L.-F., & Brodney, M. A. (2008). Therapeutic approaches for the treatment of Alzheimer's disease: an overview. *Topics in Medicinal Chemistry, 2*, 1-24.

Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: on-road versus simulated driving. *Accident Analysis and Prevention, 35*, 797-803.

Lee, H. C., & Lee, A. H. (2005). Identifying older drivers at risk of traffic violations by using a driving simulator: a 3-year longitudinal study. *The American Journal of Occupational Therapy, 59*, 97-100.

Lee, H. C., Lee, A. H., & Cameron, D. (2003). Validation of a driving simulator by measuring the visual attention skill of older adult drivers. *The American Journal of Occupational Therapy, 57*, 324-328.

Lee, H. C., Lee, A. H., Cameron, D., & Li-Tsang, C. (2003). Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. *Journal of Safety Research, 34*, 453-459.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2001). *Neuropsychological Assessment* (4 ed.). New York: Oxford University Press.

Li, G., Braver, E. R., & Chen, L.-H. (2003). Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident Analysis and Prevention, 35*, 227-235.

Li, M., Ng, T. P., Kua, E. H., & Ko, S. M. (2006). Brief informant screening test for mild cognitive impairment and early Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders, 21*, 392-402.

Lincoln, N. B., Radford, K. A., Lee, E., & Reay, A. C. (2006). The assessment of fitness to drive in people with dementia. *International Journal of Geriatric Psychiatry, 21*, 1044-1051.

Liu, L., Miyazaki, M., & Watson, B. (1999). Norms and validity for the "DriVR" - a virtual reality driving assessment for persons with head injury. *Cyberpsychology & Behavior, 2*, 53-67.

Lövsund, P., Hedin, A., & Törnros, J. (1991). Effects on driving performance of visual field defects: a driving simulator study. *Accident Analysis and Prevention, 23,* 331-342.

Lundberg, C., & Hakamies-Blomqvist, L. (2003). Driving tests with older patients: effect of unfamiliar versus familiar vehicle. *Transportation Research Part F, 6*, 163-173.

Marottoli, R. A., Cooney, L. M., Wagner, D. R., Doucette, J., & Tinetti, M. E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine, 121*, 842-846.

Marottoli, R. A., Mendes de Leon, C. F., Glass, T. A., Williams, C. S., Cooney Jr, L. M., & Berkman, L. F. (2000). Consequences of driving cessation: decreased out-of-home activity levels. *Journal of Gerontology: Social Sciences, 55B*, S334-S340.

Marson, D., & Hebert, K. R. (2006). Functional assessment. In D. K. Attix & K. A. Welsh-Bohmer (Eds.), *Geriatric Neuropsychology: Assessment and Intervention* (158-197). New York: The Guilford Press.

McGwin, G. J., Sims, R. V., Pulley, L., & Roseman, J. M. (2000). Relations among chronic medical conditions, medications, and automobile crashes in the elderly: a population-based case-control study. *American Journal of Epidemiology, 152*, 424-431.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer’s disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. *Neurology, 34*, 939-944.

Meuleners, L. B., Harding, A., Lee, A. H., & Legge, M. (2006). Fragility and crash over-representation among older drivers in Western Australia. *Accident Analysis and Prevention, 38*, 1006-1010.

Meuser, T. M., Carr, D. B., & Ulfarsson, G. F. (2009). Motor-vehicle crash history and licensing outcomes for older drivers reported as medically impaired in Missouri. *Accident Analysis and Prevention, 41*, 246-252.

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial* Lutz, FL: Psychological Assessment Resources, Inc.

Mezuk, B., & Rebok, G. W. (2008). Social integration and social support among older adults following driving cessation. *Journal of Gerontology: Social Sciences, 63B*, S298-S303.

Ministry of Transport (2008a). *Risk on the road: Household Travel Survey v1 Nov 2008*. Wellington: Ministry of Transport.

Ministry of Transport (2008b). *Motor vehicle crashes in New Zealand 2007: statistical statement calendar year 2007*. Wellington: Ministry of Transport.

Ministry of Transport (2009). *Motor Vehicle Crashes in New Zealand 2008*. Wellington: Ministry of Transport.

Ministry of Transport (2010). *Motor vehicle crashes in New Zealand 2009: statistical statement calendar year 2009*. Wellington: Ministry of Transport.

Mohs, R. C. (1994). Alzheimer's Disease Assessment Scale (1994 Revised ed.): The Mount Sinai School of Medicine.

Mohs, R. C., Knopman, D., Petersen, R. C., Ferris, S. H., Ernesto, C., Grundman, M., Sano, M., Bieliauskas, L., Geldmacher, D., Clark, C., & Thal, L. J. (1997).Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope *Alzheimer Disease and Associated Disorders*. S13-S21.

Molloy, D. W., Alemayehu, E., & Roberts, R. (1991). Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *The American Journal of Psychiatry, 148*, 102-105.

Molloy, D. W., & Standish, T. I. M. (1997). A guide to the Standardized Mini-Mental State Examination. *International Psychogeriatrics, 9* (Suppl 1), 87-94.

Molnar, F. J., Marshall, S. C., Man-Son-Hing, M., Wilson, K. G., Byszewski, A. M., & Stiell, I. (2007). Acceptability and concurrent validity of measures to predict older driver involvement in motor vehicle crashes: an Emergency Department pilot case-control study. *Accident Analysis and Prevention, 39*, 1056-1063.

Molnar, F. J., Patel, A., Marshall, S. C., Man-Son-Hing, M., & Wilson, K. G. (2006). Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: a systematic review. *Journal of the American Geriatrics Society, 54*, 1809-1824.

Monash University Accident Research Centre (2004). *Model licence re-assessment procedure for older drivers: stage 2 research*. Canberra: Austroads Incorporated.

Mungas, D. (1991). In-office mental status testing: a practical guide. *Geriatrics, 46,* 54-58.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*, 695-699.

National Highway Traffic Safety Administration (1999). *Safe mobility for older people notebook*: U.S. Department of Transportation.

New Zealand Police (2008). *Annual Report for the year ended 30 June 2008*. Wellington.

New Zealand Transport Agency (2009). *Medical aspects of fitness to drive: a guide for medical practitioners*. Palmerston North: New Zealand Transport Agency.

O'Connor, M. L., Edwards, J. D., Wadley, V. G., & Crowe, M. (2010). Changes in mobility among older adults with psychometrically defined mild cognitive impairment. *Journal of Gerontology: Psychological Sciences, 65B*, 306-316.

Organisation for Economic Co-operation and Development (2001). *Ageing and transport: mobility and safety issues*. Paris: OECD Publications.

Ott, B. R., Heindel, W. C., Papandonatos, G. D., Festa, E. K., Davis, J. D., Daniello, L. A., & Morris, J. C. (2008). A longitudinal study of drivers with Alzheimer disease. *Neurology, 70*, 1171-1178.

Owsley, C., Ball, K. K., McGwin, G., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association, 279*, 1083-1088.

Owsley, C., Ball, K. K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6,* 403-415.

Owsley, C., Stalvey, B., Wells, J. E., & Sloane, M. E. (1999). Older drivers and cataract: driving habits and crash risk. *The Journals of Gerontology: Series A: Biological Sciences and Medical Sciences, 54A*, M203-M211.

Oxley, J., Charlton, J., Scully, J., & Koppel, S. (2010). Older female drivers: an emerging transport safety and mobility issue in Australia. *Accident Analysis and Prevention, 42*, 515-522.

Parker, D., McDonald, L., Rabbitt, P., & Sutcliffe, P. (2000). Elderly drivers and their accidents: the Aging Driver Questionnaire. *Accident Analysis and Prevention, 32*, 751-759.

Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38*, 1036-1048.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*, 183-194.

Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L., & DeKosky, S. T. (2001). Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, *56*, 1133-1142.

Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., Burke, J. R., Hurd, M. D., Potter, W. L., Steffens, D. C., Willis, R. J., & Wallace, R. B. (2007). Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*, *29*, 125-132.

Ragland, D. R., Satariano, W. A., & MacLeod, K. E. (2005). Driving cessation and increased depressive symptoms. *Journal of Gerontology: Medical Sciences, 60A*, 399-403.

Rajalin, S. (1994). The connection between risky driving and involvement in fatal accidents. *Accident Analysis and Prevention, 26*, 555-562.

Rapoport, M. J., Herrmann, N., Molnar, F., Rochon, P. A., Juurlink, D. N., Zagorski, B., Seitz, D., Morris, J. C., & Redelmeier, D. A. (2008). Psychotropic medications and motor vehicle collisions in patients with dementia. *Journal of the American Geriatrics Society*, *56*, 1968-1970.

Raz, N., Rodrigue, K. M., Kennedy, K. M., & Acker, J. D. (2007). Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. *Neuropsychology*, *21*, 149-157.

Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: a real distinction? *Ergonomics*, *33*, 1315-1332.

Reger, M. A., Welsh, R. K., Watson, G. S., Cholerton, B., Baker, L. D., & Craft, S. (2004). The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. *Neuropsychology, 18*, 85-93.

Reisberg, B., Finkel, S., Overall, J., Schmidt-Gollas, N., Kanowski, S., Lehfeld, H., Hulla, F., Sclan, S. G., Wilms, H. U., Heininger, K., Hindmarch, I., Stemmler, M., Poon, L., Kluger, A., Cooler, C., Bergener, M., Hugonot-Diener, L., Robert, P. H., Antipolis, S., & Erzigkeit, H. (2001). The Alzheimer's Disease Activities of Daily Living International Scale (ADL-IS). *International Psychogeriatrics*, *13*, 163-181.

Ritchie, K., & Kildea, D. (1995). Is senile dementia "age-related" or "ageing-related"? Evidence from meta-analysis of dementia prevalence in the oldest old. *The Lancet*, *346*, 931-934.

Romanowicz, P. A., & Hagge, R. A. (1995).An evaluation of the validity of California's driving performance evaluation road test. Sacramento, CA: California Department of Motor Vehicles. *Report No. 154*,

Sagberg, F. (2006). Driver health and crash involvement: a case-control study. *Accident Analysis and Prevention, 38*, 28-34.

Schultheis, M. T., & Mourant, R. R. (2001). Virtual reality and driving: the road to better assessment for cognitively impaired populations. *Presence: Teleoperators and Virtual Environments, 10*, 431-439.

Schultheis, M. T., Rebimbas, J., Mourant, R., & Millis, S. R. (2007). Examining the usability of a virtual reality driving simulator. *Assistive Technology, 19*, 1-10.

Schultheis, M. T., Simone, L. K., Roseman, E., Nead, R., Rebimbas, J., & Mourant, R. (2006). Stopping behavior in a VR driving simulator: a new clinical measure for the assessment of driving, *28th IEEE EMBS Annual International Conference* (pp. 4921-4924). New York City.

Schwebel, D. C., Ball, K. K., Severson, J., Barton, B. K., Rizzo, M., & Viamonte, S. M. (2007). Individual difference factors in risky driving among older adults. *Journal of Safety Research, 38*, 501-509.

Schwebel, D. C., Severson, J., Ball, K. K., & Rizzo, M. (2006). Individual difference factors in risky driving: the roles of anger/hostility, conscientiousness, and sensation-seeking. *Accident Analysis and Prevention, 38*, 801-810.

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. *Traffic Injury Prevention, 10*, 379-385.

Silva, M. T., Laks, J., & Engelhardt, E. (2009). Neuropsychological tests and driving in dementia: a review of the recent literature. *Revista da Associacao Medica Brasileira, 55*, 484-488.

Sims, R. V., Ahmed, A., Sawyer, P., & Allman, R. M. (2007). Self-reported health and driving cessation in community-dwelling older drivers. *The Journals of Gerontology: Health module, 62A*, 789-793.

Sims, R. V., McGwin, G., Allman, R. M., Ball, K. K., & Owsley, C. (2000). Exploratory study of incident vehicle crashes among older drivers. *Journal of Gerontology, 55A*, M22-M27.

Sims, R. V., Owsley, C., Allman, R. M., Ball, K. K., & Smoot, T. M. (1998). A preliminary assessment of the medical and functional factors associated with vehicle crashes by older adults *Journal of the American Geriatrics Society, 46*, 556-561.

Smith, T., Gildeh, N., & Holmes, C. (2007). The Montreal Cognitive Assessment: validity and utility in a memory clinic setting. *The Canadian Journal of Psychiatry, 52*, 329-332.

Snellgrove, C. A. (2000).Cognitive screening for the safe driving competence of older people with mild cognitive impairment or early dementia. Canberra: Australian Transport Safety Bureau.

Sommer, M., Herle, M., Häusler, J., Risser, R., Schützhofer, B., & Chaloupka, C. (2008). Cognitive and personality determinants of fitness to drive. *Transportation Research Part F, 11*, 362-375.

Staplin, L., Lococo, K. H., McKnight, A. J., McKnight, A. S., & Odenheimer, G. L. (1998).Intersection negotiation problems of older drivers. Volume II: background synthesis on age and intersection driving difficulties. Washington, DC: Office of Research and Traffic Records.

Stav, W. B., Justiss, M. D., McCarthy, D. P., Mann, W. C., & Lanford, D. N. (2008). Predictability of clinical assessments for driving performance. *Journal of Safety Research, 39*, 1-7.

Stelmach, G. E., & Nahom, A. (1992). Cognitive-motor abilities of the elderly driver. *Human Factors, 34*, 53-65.

Steyerberg, E. W., Bleeker, S. E., Moll, H. A., Grobbee, D. E., & Moons, K. G. M. (2003). Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *Journal of Epidemiology, 56*, 441-447.

Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology, 54*, 774-781.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary* (3rd ed.). New York: Oxford University Press, Inc.

Sullivan, C. (2004). *Older driver crash statistics*. Wellington: Capital Research Ltd.

Sullman, M. J. M., & Baas, P. H. (2004). Mobile phone use amongst New Zealand drivers. *Transportation Research Part F, 7*, 95-105.

Sullman, M. J. M., Meadows, M. L., & Pajo, K. B. (2002). Aberrant driving behaviours amongst New Zealand truck drivers. *Transportation Research Part F, 5*, 217-232.

Systems Technology, I. (2010). Retrieved 30/01/2010, from http://www.stisimdrive.com/our-solutions/training-solutions/elderly-drivers

Tabachnick, B.G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). New York: HarperCollins.

Tefft, B. C. (2008). Risks older drivers pose to themselves and to other road users. *Journal of Safety Research, 39*, 577-582.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: a guidelines editorial. *Educational and Psychological Measurement, 55,* 525-534.

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology, 19*, 203-214.

Trobe, J. D., Waller, P. F., Cook-Flannagan, C. A., Teshima, S. M., & Bieliauskas, L. A. (1996). Crashes and violations among drivers with Alzheimer Disease. *Archives of Neurology, 53*, 411-416.

Uc, E. Y., Rizzo, M., Johnson, A. M., Dastrup, E., Anderson, S. W., & Dawson, J. D. (2009). Road safety in drivers with Parkinson disease. *Neurology, 73*, 2112-2119.

Wald, J. L., Liu, L., & Reil, S. (2000). Concurrent validity of a virtual reality driving assessment for persons with brain injury. *Cyberpsychology & Behavior, 3*, 643-654.

Warrington, E. K., & James, M. (1991). *The Visual Object and Space Perception Battery*. Bury St Edmunds, Suffolk: Thames Valley Test Company.

Wasielewski, P. (1984). Speed as a measure of driver risk: observed speeds versus driver and vehicle characteristics. *Accident Analysis and Prevention, 16*, 89-103.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - Third Edition*. San Antonio, Texas: Pearson Education, Inc.

Wechsler, D. (2001). *Wechsler Test of Adult Reading*. San Antonio, Texas: Pearson Education, Inc.

Whelihan, W. M., DiCarlo, M. A., & Paul, R. H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology, 20*, 217-228.

Witten, I. H., & Frank, E. (2000). *Data mining: practical machine learning tools and techniques with Java implementations*. San Francisco: Morgan Kaufmann.

Wood, J. M., Anstey, K. J., Kerr, G. K., Lacherez, P. F., & Lord, S. (2008). A multidomain approach for predicting older driver safety under in-traffic road conditions. *Journal of the American Geriatrics Society, 56*, 986-993.

Zec, R. F., Landreth, E. S., Vicari, S. K., Belman, J., Feldman, E., Andrise, A., Robbs, R., Becker, R., & Kumar, V. (1992). Alzheimer Disease Assessment Scale: a subtest analysis. *Alzheimer Disease and Associated Disorders, 6*, 164-181.

Zuin, D., Ortiz, J., & Lopez, O. L. (2002). Motor vehicle crashes and abnormal driving behaviours in patients with dementia in Mendoza, Argentina. *European Journal of Neurology*, *9*, 29-34.

CHAPTER 11 –

Appendices

11.1 Appendix A – Published Paper 1

Accident Analysis & Prevention
Volume 42, Issue 6, November 2010, Pages 1759-1768

doi:10.1016/j.aap.2010.04.017 | How to Cite or Link Using DOI
Permissions & Reprints

Comparison of a linear and a non-linear model for using sensory–motor, cognitive, personality, and demographic data to predict driving ability in healthy older adults

Petra A. Hoggarth[a, b], **Carrie R.H. Innes**[a, c], **John C. Dalrymple-Alford**[a, b, d], **Julie E. Severinsen**[e] and **Richard D. Jones**[a, b, c, d, f]

[a] Van der Veer Institute for Parkinson's and Brain Research, 66 Stewart Street, Christchurch 8011, New Zealand

[b] Department of Psychology, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

[c] Department of Medical Physics and Bioengineering, Christchurch Hospital, Private Bag 4710, Christchurch, New Zealand

[d] Department of Medicine, University of Otago, PO Box 4345, Christchurch 8140, New Zealand

[e] Department of Occupational Therapy, Burwood Hospital, Private Bag 4708, Christchurch, New Zealand

[f] Department of Electrical and Computer Engineering, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Received 12 August 2009;
revised 8 March 2010;
accepted 22 April 2010.
Available online 26 May 2010.

Abstract

This study compared the ability of binary logistic regression (BLR) and non-linear causal resource analysis (NCRA) to utilize a range of cognitive, sensory–motor, personality and demographic measures to predict driving ability in a sample of cognitively healthy older drivers.

Participants were sixty drivers aged 70 and above (mean = 76.7 years, 50% men) with no diagnosed neurological disorder. Test data was used to build classification models for a Pass or Fail score on an on-road driving assessment. The generalizability of the models was estimated using leave-one-out cross-validation.

Sixteen participants (27%) received an on-road Fail score. Area under the ROC curve values were .76 for BLR and .88 for NCRA (no significant difference, $z = 1.488$, $p = .137$). The ROC curve was used to select three different cut-points for each model and to compare classification. At the cut-point corresponding to the maximum average of sensitivity and specificity, the BLR model had a sensitivity of 68.8% and specificity of 75.0% while NCRA had a sensitivity of 75.0% and specificity of 95.5%. However, leave-one-out cross-validation reduced sensitivity in both models and particularly reduced specificity for NCRA.

11.2 Appendix B – Published Paper 2

DOI: 10.1111/j.1741-6612.2010.00478.x

Policy Update

Assessment of older drivers in New Zealand: The current system, research and recommendations

Petra Hoggarth
Van der Veer Institute for Parkinson's and Brain Research, Christchurch; and Department of Psychology, University of Canterbury, Christchurch, New Zealand

Carrie Innes
Van der Veer Institute for Parkinson's and Brain Research, Christchurch; and Department of Medical Physics and Bioengineering, Canterbury District Health Board, Christchurch, New Zealand

John Dalrymple-Alford
Van der Veer Institute for Parkinson's and Brain Research, Christchurch; and Department of Psychology, University of Canterbury, Christchurch; and Department of Medicine, University of Otago, Christchurch, New Zealand

Matthew Croucher
Psychiatry Service for the Elderly, Canterbury District Health Board, Christchurch; and Department of Psychological Medicine, University of Otago, Christchurch, New Zealand

Julie Severinsen and Jane Gray
Department of Occupational Therapy, Canterbury District Health Board, Christchurch, New Zealand

Jenny Oxley and Barbara Brook
OTRS Group Ltd, Hamilton, New Zealand

Paul Abernethy
Pegasus Health Ltd, Christchurch, New Zealand

Richard Jones
Van der Veer Institute for Parkinson's and Brain Research, Christchurch; Department of Psychology, University of Canterbury, Christchurch; Department of Medical Physics and Bioengineering, Canterbury District Health Board, Christchurch; Department of Medicine, University of Otago, Christchurch; and Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand

To clarify the current New Zealand driving licensing requirements for older adults and to provide practical recommendations for those health professionals who make decisions regarding driving ability in older adults. Health professionals involved in the assessment of older drivers were asked to clarify areas where more efficient use could be made of assessment resources. A review of driving literature was performed to find specific factors associated with increased risk of negative driving outcomes in older adults. Particular focus was made of the suitability of different types on on-road assessment for certain patient groups, the effect of specific diseases and medications on driving safety, and the effect of cognitive impairment. A list of seven recommendations were compiled which include a focus on appropriate on-road driving assessment referral, driver refresher courses, cognitive screening for those presenting for licence renewal and sensitive broaching of the topic of driving cessation.

Key words: *aged, automobile driving, health planning, licensure, risk assessment.*

Correspondence to: •• Petra Hoggarth, Van der Veer Institute for Parkinson's and Brain Research. Email: petra.hoggarth@vanderveer.org.nz

Introduction

This article is a collaborative effort between researchers, driving specialist occupational therapists (OTs), an old-age psychiatrist and a general practitioner (GP) representative to clarify the current older driver relicensing requirements in New Zealand, to provide accurate summaries of international research investigating risk factors for unsafe driving in older people, and to provide concrete recommendations for how GPs and other health professionals charged with assessing driving ability can make the best use of time and resources to detect those older drivers who are at increased risk for unsafe driving.

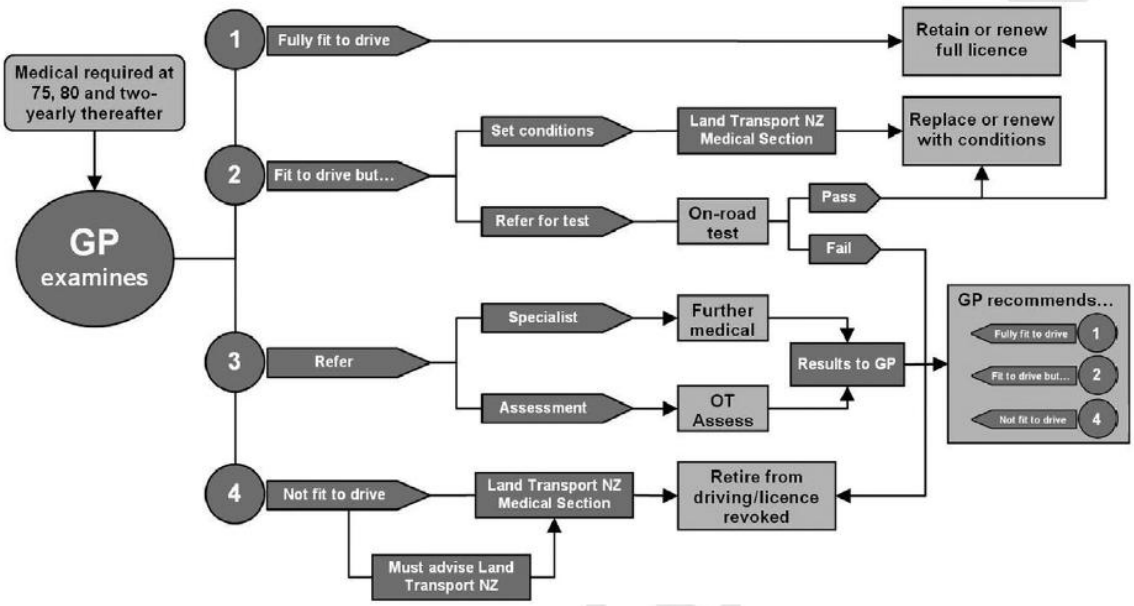
Driving assessment in New Zealand

Up until December 2006, New Zealand drivers aged 80 and over were required to sit biennial on-road driving assessments in order to retain their drivers licence. This policy was ended in 2006 because of accusations that it was ageist. Drivers are now required to obtain a 'medical fitness to drive' certificate from their GPs at ages 75, 80 and biennially thereafter. A guide exists for medical practitioners, optometrists and OTs for assessing medical fitness to drive [1]. The guide is not specific to older drivers but has sections on medical illnesses more common in older people. Fact sheets related to older driver relicensing are available on the New Zealand Transport Agency's (NZTA; formerly Land Transport New Zealand) website (http://www.landtransport.govt.nz/information-for/older-drivers.html). The NZTA provided a flow chart for GPs in 2006 detailing the steps involved in older driver licence renewal (Figure 1).

The chart depicts four decision pathways. The most direct options (1 and 4) are for a GP to decide independently whether a patient is medically fit to drive and to recommend this unconditionally. These options should be used when the GP is satisfied that their patient is fit to drive (Option 1), or when there is obvious impairment such as moderate to severe

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA
UNCORRECTED PROOF

Figure 1: Flowchart of the older driver licence renewal system (adapted from NZTA information pack for GPs of December 2006). NZTA, New Zealand Transport Agency; GP, general practitioner.



dementia or levels of visual acuity that do not meet NZTA minimum requirements (Option 4). Option 2 has two sub-options. The first is for the GP to supply a medical fitness to drive certificate with set conditions added to the licence, such as must wear corrective lenses, distance restrictions or daytime-only restrictions, although there is no evidence that these latter two conditions reduce on-road errors and crashes.

The second sub-option is for the GP to supply a medical fitness to drive certificate subject to the patient satisfactorily undertaking an On-road Safety Test. If the patient fails that test, this information is forwarded to NZTA who make a decision regarding renewal of the patient's licence.

Option 3 is an intermediate step for when the GP considers a diagnosed or suspected medical illness may be affecting driving safety. In this case, a GP can refer their patient to a medical specialist (e.g. neurologist, geriatrician, psychiatrist, optometrist), or to an OT who can perform a Medical Driving Assessment. On receipt of the specialist assessment results, the GP follows flowchart options 1, 2 or 4. The NZTA state in their GP information pack (August 2006) that a patient who has received a favourable report following a Medical Driving Assessment performed by an OT should not be referred for an On-road Safety Test because the Medical Driving Assessment is a more thorough investigation of a person's ability to drive. If a GP decides that their patient is not fit to drive, they should inform the patient and ask them to cease driving immediately. Only a medical practitioner is able to sign a medical certificate (DL9), or a registered optometrist to sign an eyesight certificate (DL12), stating that an appropriate examination has been undertaken and that a patient is considered medically fit to renew their licence. Specialist assessors and OTs may also supply recommendations directly to the NZTA, who ultimately make the legal decision to retain or remove driving privileges.

General practitioners must determine not only a patient's medical fitness to drive but also their ability to drive safely. The LTNZ website states: 'If your doctor decides you are medically fit to drive, but is unsure about your ability to drive safely, he or she may issue you with a medical certificate for driver licence enabling you to renew your driver licence, provided you pass an On-road Safety Test with a testing officer' [2]. Medical fitness to drive indicates either that there is no medical disorder or that if there is, and if it has the potential to affect driving, it is not doing so in a particular individual. Ability to drive safely includes a person's knowledge and on-road application of road rules and competent driving behaviour, and can be independent of a medical condition. We contend that determining ability to drive safely can only be performed following the observation of actual driving behaviour and is, thus, almost impossible to do in a GP's clinic, unless there is reliable information attesting to a person exhibiting unsafe driving behaviour. Determining ability to drive safely may be one of the hardest decisions GPs are asked to make regarding driving ability in older drivers, particularly because a person does not have to be medically unwell to display unsafe driving behaviour.

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

The On-road Safety Test and the Medical Driving Assessment are discussed below (specialist assessments by optometrists are not considered in this article). Suggestions for selecting which assessment to refer a patient to are provided in the recommendations section of this article.

The On-road Safety Test is a 30-minute on-road assessment essentially identical to the previous compulsory Older Driver Test. On-road Safety Tests are provided by the New Zealand Automobile Association, Vehicle Testing New Zealand and Vehicle Inspection New Zealand (http://www.ltsa.govt.nz/licensing/older/guide-on-road-safety-test.html). The test assesses basic driving skills (e.g. leaving the kerb, turning left at an intersection), hazard detection (e.g. negotiating a crossroad, stopping or giving way at Stop or Give Way signs) and more complex driving situations (e.g. turning right at a crossroad in medium-to-heavy traffic in a 50 km/hour zone). Scoring is based on the performance of predetermined manoeuvres, with error scores weighted and combined to give an ultimate pass or fail score. This test is less comprehensive and shorter in duration than the New Zealand full-licence practical test for novice drivers.

Medical Driving Assessments are performed by OTs with specialist training in driver assessment. There are few medical conditions which are automatic rule-outs for driving licensure. Medical illnesses that affect driving usually do so unpredictably. Medical Driving Assessments do not make use of predetermined lists of errors with weighted scores, unlike the On-road Safety Test, but are considered by NZTA to be a more comprehensive assessment of driving ability than the On-road Safety Test. A Medical Driving Assessment fulfils requirements for assessing both a patient's medical fitness to drive and their ability to drive safely. The on-road component of an assessment averages around 45 minutes in length and spans a wide range of on-road driving situations (e.g. different speed zones, single- and multi-laned roads, moving from quieter to more busy roads). OTs assess aspects such as awareness of the road and traffic environment, apparent insight into the driving task and any errors that occur, and whether patients are able to compensate for difficulties posed by their medical condition. Stringent observance of all road rules is not the primary focus; people are asked to drive as they normally would. The objective is not to penalise people for driving errors commonly performed by many drivers, but rather to determine whether a medical illness has made the person unable to drive safely. An OT assessment of driver safety is based on a combination of outcomes of on-road assessment, cognitive and physical evaluations, and any information gathered from a patient including, but not limited to, driving history and frequency, self-imposed limitations, and orientation to time and place. A specialist OT assessment includes consideration of whether driving instruction or modification to a vehicle could improve performance to safe standards; this is not provided by the On-road Safety Test.

Before the abolition of compulsory on-road assessment for older drivers in December 2006, a New Zealand study investigated the association between the outcome of the compulsory Older Driver Test and involvement in a subsequent fatal or injury-causing crash over a 2-year period in over 39 000 drivers aged 80 and above [3]. The previous testing system stipulated no limit to the number of times drivers could resit the compulsory test in order to secure a pass. Seventy-eight percent of the study sample passed the assessment on the first administration, and 17% passed following two or more tests (two drivers passed the test on the 10th resit). The remaining 5% of drivers took at least one test and never received a passing grade. The number of drivers who decided to relinquish their licence rather than submit to a test is unknown.

Keall and Frith [3] found that for every time the older Driver Test was repeated in order to secure a pass the risk of being involved in an injury-causing crash increased by 33%. In reality, however, the number of people involved in serious crashes was small in both groups – 223 in 32 135 first-time passers, versus 80 in 6863 who passed after two or more attempts. Presumably, the crash rate of those who never passed the on-road assessment would have been higher than in these two groups in which a pass was obtained, but their crash rate is unknown as a fail outcome equated to a loss of licence and supposed driving abstinence. This difficulty applies to all studies of on-road driving assessments and is a major obstacle in assessing their reliability and validity. Obviously allowing drivers considered 'unsafe' to continue driving in order to monitor their performance has ethical implications. Notwithstanding, we have performed a 2-year follow-up of 60 older drivers who failed a single on-road driving assessment and who were allowed to continue driving (participants had no neurological diagnosis and had not been referred for assessment). Results at 1-year follow-up showed no trends for higher self- or police-reported crashes or traffic offences in the fail group compared with the pass group (15 failed and 43 passed) [4].

An impediment to the access of on-road driving assessments in New Zealand is that only a small number are publicly funded in only a few regions. An OT-administered driving assessment performed privately can cost as much as $NZ600. In contrast, an On-road Safety Test costs only $NZ41 (following a $NZ7.90 government subsidy), although it cannot evaluate the effects of medical conditions on driving behaviour. The substantial cost of private specialist driving assessment may dissuade GPs from referring their older patients for a Medical Driving Assessment. Thus, increased availability of publicly funded Medical Driving Assessments deserves consideration by health boards. Also, the current system for dealing with a Fail on the On-road Safety Test as depicted in Figure 1 is for NZTA to revoke a person's licence. As cited above, 17% of older drivers who initially failed the older Driver Test passed on subsequent sittings – albeit with a small increase in the number of serious crash incidences in those who repeated the test [3]. We there-

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

UNCORRECTED PROOF

fore suggest that GPs strongly encourage their patients to attend a driver refresher course prior to taking the On-road Safety Test in order to maximise the likelihood that safe drivers will pass the test on a single administration. Details on driver refresher courses are contained in the recommendations section below.

Risk factors for decreased driving safety of older drivers

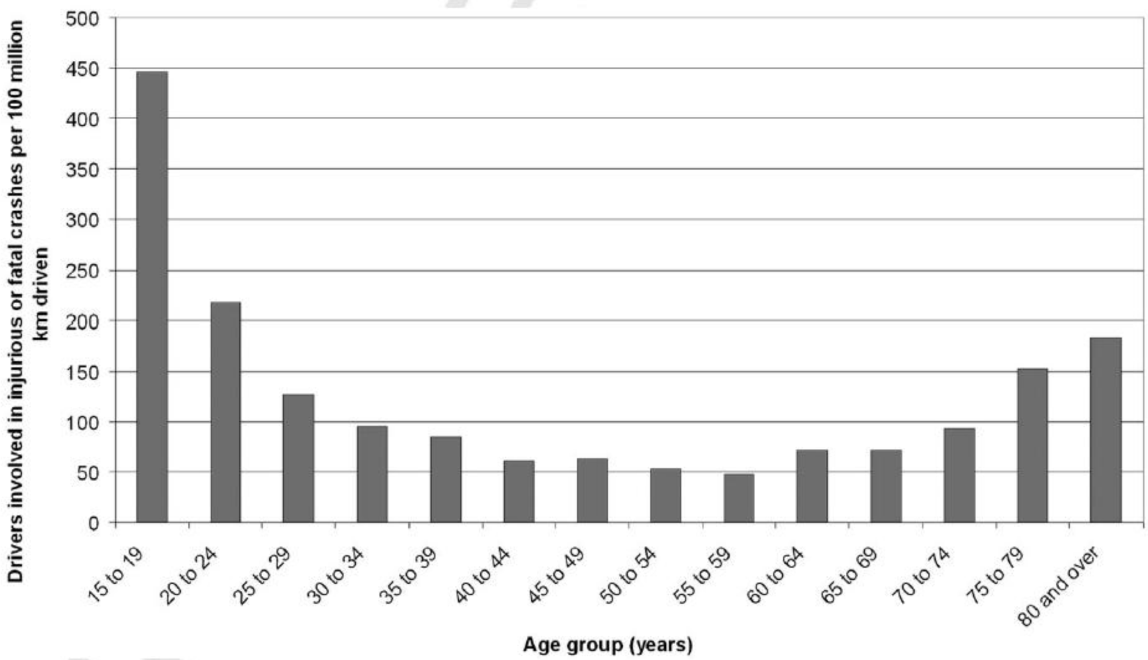
By distance driven, drivers aged 60 and above have increased injury and death rates compared with middle-aged drivers (Figure 2) [5–8]. A higher burden of traffic-related morbidity in older population is therefore expected in future because of the increasing proportion of older people in society in the coming decades.

Older drivers are more physically fragile. Their death and injury rates are thus inflated by accidents that would not lead to injury or death in younger drivers [9–11]. Annual distance driven in older age groups is substantially lower compared with middle-aged drivers. These biases are minimised when it is shown that older adults' higher injury and death rates persist when statistics are adjusted for the number of driving trips and for increased physical fragility [5].

Factors linked to unsafe driving in older adults include heart disease, cerebrovascular disease, poor visual attention and mental flexibility, age-related cognitive decline, and dementia [11–18]. Medications associated with poorer driving or crashes include hypnotics, non-steroidal anti-inflammatories, angiotensin converting enzyme inhibitors, anticoagulants, benzodiazepines, tricyclic antidepressants and lithium [17,19–22]. The effect of haloperidol on driving has not been specifically investigated in older adults but it has been associated with poorer driving-related psychomotor performance in younger drivers. Compared with haloperidol, atypical antipsychotics are associated with fewer psychomotor problems [23,24]. Although there have been few studies, sodium valproate and carbamazepine have not been shown to have a consistent association with increased crash risk [21,25]. A recent review of older drivers concluded that tricylic antidepressant doses should be introduced gradually, their negative effects on driving being most likely to occur in the first week of taking a tricylic, and that more severe depressive symptoms may also negatively affect driving [19]. Unfortunately, the interaction effects of multiple medications on driving safety, especially in older adults, has not been studied. Antidepressants prescribed along with sedating medications such as benzodiazepines may have an adverse synergistic effect. For example, fluoxetine prescribed to a patient already on alprazolam may increase benzodiazepine side effects such as psychomotor impairment through effects on the cytochrome P-450 2D6 enzyme [19]. The Medical Aspects of Fitness to Drive guidebook has lists of medications that should be considered in relation to medical fitness to drive [1].

As a group, drivers with dementia have crash rates 2.5 to 10.7 times higher than those without dementia [14,26].

Figure 2: Number of New Zealand drivers per age group involved in injury and fatality crashes per 100 million km driven. Data obtained from the Household Travel Survey [6].



© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

Studies of dementia prevalence have reported rates between 13% and 43% in the 80 to 89 age group, increasing exponentially per year within this age range, and between 40% to 65% in those aged over 90 [27–29]. Many people with early dementia may pass an on-road driving assessment [30] but there are as yet no reliable neuropsychological or demographic variables that can discriminate between individuals who pass and fail. Even if a driver with dementia is judged to be safe, driver reassessment will need to recur for as long as the patient continues to drive. Fox et al. [31] found that of seven drivers with dementia who passed an on-road driving assessment, only three passed a second assessment 6 months later. Other researchers have also suggested that drivers with dementia should be assessed on-road every 6 months [32,33].

In New Zealand, a computerised test battery of Sensory-motor and Cognitive Tests (*SMCTests*™ – http://www.neurotech.org.nz/files/CanDAT_SMCTests_User_Manual.pdf) [34,35] is currently being used in several occupational therapy settings as part of comprehensive assessment of people with brain disorders [35]. A study is underway to determine whether a subset of tests in this computerised battery, or standard neuropsychological tests, can detect which drivers with Alzheimer's type dementia or mild cognitive impairment will fail an on-road driving assessment. However, no current off-road driving assessment methods have high enough levels of sensitivity and specificity for detecting unsafe driving in order to be relied on solely [36].

Driving is important for older people, as cessation is related to decreases in social participation [37,38], decreases in physical health [37], increases in depression [39] and increased mortality [40]. Although drivers aged 60 and above have increased rates of serious injury and death on the road per km driven, Evans [41] reported that per-person per-year a 70-year-old man driver poses 40% less threat to other road users than a 40-year-old man driver. A 20-year-old man driver poses a 196% greater threat to other road users than that posed by a 70-year-old man driver. This is due to the considerably lower mileage driven by older drivers which reduces their overall exposure to the risk of a crash. Because of the negative consequences of driving cessation for older adults, it is imperative that people are able to retain their licences for as long as they are able to drive safely.

The most difficult step to take in determining driver fitness is focusing on an individual's deficits and abilities, which are not easily identified from studies that have focused on group differences between safe and unsafe driver groups. Hence, pragmatic decisions made by expert driving assessors will continue to be part of the assessment of safe driving in older adults.

Recommendations for making decisions about driving for older adults

While recognising regional difficulties in availability of specialist driving services, we propose several recommendations to health practitioners to assist their decisions about driving for older adult patients. These suggestions are based on the driving research literature, our professional opinion and the currently required procedure for renewing the licences of older New Zealand drivers.

Choosing an on-road assessment

The two on-road driving assessments available (On-road Safety Test and Medical Driving Assessment) are designed for different patient groups. If a general medical condition has been diagnosed and the GP has concerns that the condition may affect driving safety, a Medical Driving Assessment is recommended, particularly if the medical condition is neurological or neuropsychiatric (e.g. dementia). Alternatively, if a patient has no diagnosed medical condition thought likely to affect driving and yet the GP is seriously concerned about their ability to drive safely, they should be referred for an On-road Safety Test.

Not everyone with a neuropsychiatric condition requires a Medical Driving Assessment. For example, nearly everyone with a stable anxiety or depressive disorder is capable of continuing to drive safely. When medical contributions to fitness to drive have made a significant effect, it is important to treat and improve these conditions as much as possible, and OTs who perform Medical Driving Assessments may be able to offer remediation and recommendations to assist a person to regain a safe driving status.

Driver refresher courses and the On-road Safety Test

To provide the best chance that an older driver will pass the On-road Safety Test, a GP should strongly encourage a driving refresher course prior to taking the test. A recent review article concluded that education interventions combined with an on-road driving component are successful in increasing driving knowledge and driving-specific skills in older people [42]. Unfortunately, the New Zealand Safe with Age older driver education course was cancelled in August 2009. While a new older driver education programme is being planned for implementation over the next few years, there is currently no publicly funded course available. The Automobile Association and other private driving schools provide older driver refresher courses.

Cognitive impairment and driving

Because of the high prevalence of cognitive impairment in the age groups for whom mandatory driving fitness certification is required, GPs should screen their patients for cognitive decline when they present for a medical fitness to drive certificate. Illnesses associated with cognitive impairment include the various dementias, Parkinson's disease, multiple sclerosis, stroke and depression. An initial evaluation seeks evidence for a recent history of cognitive decline from the patient or preferably a reliable informant, observing cognitive function during the interview, and also using a formal screening test. Patients screened as positive for cognitive impairment require a thorough diagnostic evaluation. Fol-

UNCORRECTED PROOF

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

lowing diagnosis, a global management plan should be constructed that includes, but is not limited to, driving.

Patients with diagnosed progressive dementia may be, or will at some point become, unsafe drivers. In general, few people with moderate to severe dementia pass on-road assessments, but as many as half of drivers with very mild to mild dementia can pass an on-road assessment [30,43]. At least six monthly assessments are reasonable for those who continue to drive – the compulsory 2-year Medical Driving Assessments are too infrequent for monitoring the effect of dementia on driving. Driving fitness certificates can be issued for this shorter time to ensure reviews continue. The threshold for requesting on-road or specialist assessments should be low. Discussions regarding cessation of driving and alternative arrangements for transportation are best had when the patient is still capable of making effective decisions. A caregiver's rating of a patient's driving ability as marginal or unsafe has been shown to be related to adverse on-road outcomes, but patient's self-ratings of driving ability are not [43]. Fitness to drive should also be assessed after any introduction or increase in dose of medications that may increase risk in people with cognitive impairment. With respect to dementia, a recent review provided an algorithm for evaluating the risk of adverse driving events, based on a Clinical Dementia Rating and driving behaviour questionnaires [43].

Formal cognitive testing and driving

The NZTA offers a short test of road sign recognition in the Medical Aspects of Fitness to Drive [1] manual to use with people who may be suspected of having dementia. As there has been no research to assess the validity of this test, GPs are advised instead to use standardised and validated assessment measures to identify cognitive impairment coupled with pragmatic history-taking for driving ability. A popular screening test for significant cognitive impairment is the Mini Mental State Exam (MMSE) [44]. GPs should be aware that this test is not sensitive enough to detect mild cognitive impairment, but obtaining a score of 25/30 or below on the MMSE may indicate the presence of cognitive impairment such as from a dementia (the test is more specific than it is sensitive) [45,46]. A standardised version of the MMSE (the SMMSE) [47,48] is available at no cost from the Van der Veer Institute for Parkinson's and Brain Research (http://www.vanderveer.org.nz/files/SMMSE.pdf). The original article provides further guidance on administration [48]. A more sensitive screen for mild cognitive impairment and mild dementia is the Montreal Cognitive Assessment [49], available online (http://www.mocatest.org/).

Driving risk history assessment

The absence of recent crashes and traffic offences does not necessarily mean that a person is a safe driver. Crashes are low base-rate occurrences, even for impaired drivers, and other road users are often able to avoid unsafe drivers. Inquiring about 'discussions' with police officers, whether or not an infringement notice was administered, or whether other drivers have been complaining about the driver's behaviour may provide useful information. Assessing the car for damage can also be informative. In a document titled 'Supporting older drivers' [50], the NZTA specifically advises family members to contact a person's doctor if they are concerned about safety risks. Reports of deterioration in driving ability can be elicited, although family members are sometimes reluctant to talk about the issue, particularly in the presence of the affected family member. It is important to remember that holding a driver licence is a privilege and not a right. The Privacy Act 1993 states that 'disclosure of personal information may be necessary to prevent or lessen a serious and imminent threat to: (i) public health or public safety or (ii) the life and health of the individual concerned or another individual'. This enables GPs to investigate potentially serious driving problems.

Vision assessment

It is also helpful that drivers referred for driving assessments be first screened for meeting NZTA minimum requirements for visual field (140°) and visual acuity (6/12 binocular) [1]. Referral to an optometrist should be considered if a patient does not meet these standards.

Driving cessation

When a person must immediately cease driving, common sense actions and the assistance of the police may be required. Family or supporters will need to take responsibility for access to car keys, disabling or removing a vehicle in some situations. When assessment of driving is delayed, driving may need to cease in the interim. Many people benefit from making a ritual out of making the important decision to cease driving. Facilitating patients with impaired driving ability to surrender their licences at their local police station can be a helpful step in the grieving of one's ability to drive. Informal management of driving cessation is preferred but the threshold for formal reporting to the NZTA and/or police, depending upon urgency, must be low. Assessing patients 6 months post driving cessation to monitor the use of alternate transport is warranted. The health of older people (mental and physical) is threatened by social isolation and many New Zealanders live in areas in which social engagement is predicated upon an ability to travel.

Acknowledgements

This article has received no specific financial support. It was written during the lead author's PhD project on driving ability in healthy older adults and persons with Alzheimer's dementia. The empirical studies for the PhD project received funding from the Transport Research and Educational Trust Board, Accident Compensation Corporation, Christchurch Neurotechnology Research Programme and University of Canterbury.

UNCORRECTED PROOF

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

Key Points

- A range of physical and cognitive disorders as well as medications have been associated with increased risk of unsafe driving in older drivers.
- General practitioners and health professionals should actively screen for cognitive impairment in older drivers presenting for licence renewal.
- On-road driving assessments should only be requested for those older drivers who present with risk factors for unsafe driving, or for whom there is reliable evidence of unsafe driving behaviour.

References

1 New Zealand Transport Agency. *Medical Aspects of Fitness to Drive: A Guide for Medical Practitioners.* Palmerston North: New Zealand Transport Agency, 2009.
2 New Zealand Transport Agency. *Renewing Your Driver Licence at Age 75 and over 2009.* [Cited 9 September 2009.] Available from URL: http://www.ltsa.govt.nz/licensing/older/renewing-your-dl.html
3 Keall MD, Frith WJ. Association between older driver characteristics, on-road driving test performance, and crash liability. *Traffic Injury Prevention* 2004; 5: 112–116.
4 Hoggarth P, Jones R, Innes C, Dalrymple-Alford J. Driving assessment and subsequent driving outcome: A prospective study of safe and unsafe healthy driver groups. *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.* 2009; 433–439.
5 Tefft BC. Risks older drivers pose to themselves and to other road users. *Journal of Safety Research* 2008; 39: 577–582.
6 Ministry of Transport. *Risk on the Road: Household Travel Survey V1 Nov 2008.* Wellington: Ministry of Transport, 2008.
7 Organisation for Economic Co-operation and Development. *Ageing and Transport: Mobility and Safety Issues.* Paris: OECD Publications, 2001.
8 Langford J, Koppel S. The case for and against mandatory age-based assessment of older drivers. *Transportation Research Part F* 2006; 9: 353–362.
9 Li G, Braver ER, Chen L-H. Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident Analysis and Prevention* 2003; 35: 227–235.
10 Sullivan C. *Older Driver Crash Statistics.* Wellington: Capital Research Ltd, 2004.
11 Meuleners LB, Harding A, Lee AH, Legge M. Fragility and crash over-representation among older drivers in Western Australia. *Accident Analysis and Prevention* 2006; 38: 1006–1010.
12 Ball KK, Owsley C. Identifying correlates of accident involvement for the older driver. *Human Factors* 1991; 33: 583–595.
13 Ball KK, Owsley C. The useful field of view test: A new technique for evaluating age-related declines in visual function. *Journal of the American Optometric Association* 1993; 64: 71–79.
14 Cooper PJ, Tallman K, Tuokko H, Beattie BL. Vehicle crash involvement and cognitive deficit in older drivers. *Journal of Safety Research* 1993; 24: 9–17.
15 Dobbs AR, Heller RB, Schopflocher D. A comparative approach to identify unsafe older drivers. *Accident Analysis and Prevention* 1998; 30: 363–370.
16 Johansson K, Bronge L, Lungberg C et al. Can a physician recognize an older driver with increased crash risk potential? *Journal of the American Geriatrics Society* 1996; 44: 1198–1204.
17 McGwin GJ, Sims RV, Pulley L, Roseman JM. Relations among chronic medical conditions, medications, and automobile crashes in the elderly: A population-based case-control study. *American Journal of Epidemiology* 2000; 152: 424–431.
18 McKnight AJ, McKnight AS. Multivariate analysis of age-related driver ability and performance deficits. *Accident Analysis and Prevention* 1999; 31: 445–454.
19 Sansone RA, Sansone LA. Driving on antidepressants: Cruising for a crash? *Psychiatry (Edgemont)* 2009; 7: 13–16.
20 Leveille SG, Buchner DM, Koepsell TD et al. Psychoactive medications and injurious vehicle collisions involving older drivers. *Epidemiology* 1994; 5: 591–598.
21 Etminan M, Hemmelgarn B, Delaney JAC, Suissa S. Use of lithium and the risk of injurious motor vehicle crash in elderly adults: Case-control study nested within a cohort. *British Medical Journal* 2004; 328: 558–559.
22 Sims RV, McGwin G, Allman RM et al. Exploratory study of incident vehicle crashes among older drivers. *Journal of Gerontology* 2000; 55A: M22–M27.
23 Brunnauer A, Laux G, Zwik S. Driving simulator performance and psychomotor functions of shizophrenic patients treated with antipsychotics. *European Archives of Psychiatry and Clinical Neuroscience* 2009; 259: 483–489.
24 Kagerer S, Winter C, Möller HJ, Soyka M. Effects of haloperidol and atypical neuroleptics on psychomotor performance and driving ability in schizophrenic patients. *Neuropsychobiology* 2003; 47: 212–218.
25 Bramness JG, Skurtveit S, Neutel CI et al. An increased risk of road traffic accidents after prescriptions of lithium or valproate? *Pharmacoepidemiology and Drug Safety* 2009; 18: 492–496.
26 Zuin D, Ortiz J, Lopez OL. Motor vehicle crashes and abnormal driving behaviours in patients with dementia in Mendoza, Argentina. *European Journal of Neurology* 2002; 9: 29–34.
27 Ritchie K, Kildea D. Is senile dementia 'age-related' or 'ageing-related'? Evidence from meta-analysis of dementia prevalence in the oldest old. *Lancet* 1995; 346: 931–934.
28 Plassman BL, Langa KM, Fisher GG et al. Prevalence of dementia in the United States: The aging, demographics, and memory study. *Neuroepidemiology* 2007; 29: 125–132.
29 De Ronchi D, Berardi D, Menchetti M et al. Occurrence of cognitive impairment and dementia after the age of 60: A population-based study from Northern Italy. *Dementia and Geriatric Cognitive Disorders* 2005; 19: 97–105.
30 Berndt A, Clark M, May E. Dementia severity and on-road assessment: Briefly revisited. *Australian Journal on Ageing* 2008; 27: 157–160.
31 Fox GK, Bowden SC, Bashford GM, Smith DS. Alzheimer's disease and driving: Prediction and assessment or driving performance. *Journal of the American Geriatrics Society* 1997; 45: 949–953.
32 Adler G, Rottunda S, Dysken M. The older driver with dementia: An updated literature review. *Journal of Safety Research* 2005; 36: 399–407.
33 Duchek JM, Carr DB, Hunt L et al. Longitudinal driving performance in early-stage dementia of the Alzheimer type. *Journal of the American Geriatrics Society* 2003; 51: 1342–1347.
34 Innes CRH, Jones RD, Anderson TJ et al. Performance in normal subjects on a novel battery of driving-related sensory-motor and cognitive tests. *Behavior Research Methods* 2009; 41: 284–294.
35 Innes CRH, Jones RD, Dalrymple-Alford JC et al. Sensory-motor and cognitive tests can predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences* 2007; 260: 188–198.
36 Innes CRH, Jones R, Dalrymple-Alford J et al. Prediction of driving ability in people with dementia- and non-dementia-related brain disorders. *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.* 2009; 342–348.
37 Edwards JD, Lunsman M, Perkins M et al. Driving cessation and health trajectories in older adults. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2009; 64A: 1290–1295. 5
38 Mezuk B, Rebok GW. Social integration and social support among older adults following driving cessation. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 2008; 63B: S298–S303.
39 Ragland DR, Satariano WA, MacLeod KE. Driving cessation and increased depressive symptoms. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2005; 60A: 399–403.
40 Edwards JE, Perkins M, Ross LA, Reynolds SL. Driving status and three-year mortality among community-dwelling older adults. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2009; 64A: 300–305.
41 Evans L. Risks older drivers face themselves and threats they pose to other road users. *International Journal of Epidemiology* 2000; 29: 315–322.
42 Korner-Bitensky N, Kua A, von Zweck C, Van Benthem K. Older driver retraining: An updated systematic review of evidence of effectiveness. *Journal of Safety Research* 2009; 40: 105–111.

© 2010 The Authors
Australasian Journal on Ageing © 2010 ACOTA

43 Iverson DJ, Gronseth GS, Reger MA et al. Practice parameter update: Evaluations and management of driving risk in dementia. *Neurology* 2010; 74: 1316–1324.

44 Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinican. *Journal of Psychiatric Research* 1975; 12: 189–198.

45 Nasreddine ZS, Phillips NA, Bédirian V et al. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 2005; 53: 695–699.

46 Smith T, Gildeh N, Holmes C. The Montreal Cognitive Assessment: Validity and utility in a memory clinic setting. *The Canadian Journal of Psychiatry* 2007; 52: 329–332.

47 Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *The American Journal of Psychiatry* 1991; 148: 102–105.

48 Molloy DW, Standish TIM. A guide to the Standardized Mini-Mental State Examination. *International Psychogeriatrics* 1997; 9: 87–94.

49 Nasreddine Z. *Montreal Cognitive Assessment.* [Cited 16 October 2009.] Available from URL: http://www.mocatest.org/

50 Land Transport New Zealand. *Supporting Older Drivers: Help Your Older Friend Or Relative Stay Mobile – Safely.* Wellington: Land Transport New Zealand, 2006.

11.3 Appendix C – Published Paper 3

PROCEEDINGS of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design

DRIVING ASSESSMENT AND SUBSEQUENT DRIVING OUTCOME: A PROSPECTIVE STUDY OF SAFE AND UNSAFE HEALTHY DRIVER GROUPS

Petra Hoggarth[1,2], Richard Jones[1,3,4], Carrie Innes[1,3] & John Dalrymple-Alford[1,2]
[1]Van der Veer Institute for Parkinson's & Brain Research
[2]Dept Psychology, University of Canterbury
[3]Dept Medical Physics & Bioengineering, Christchurch Hospital,
[4]Dept Medicine, University of Otago
Christchurch, New Zealand
E-mail: petra.hoggarth@vanderveer.org.nz

Summary: Older drivers are an increasingly numerous section of the population who are often targeted for driving assessment. Little is known as to whether on-road driving assessments result in an older driver population who have fewer negative driving events. Fifty-eight healthy older drivers (mean age 77, range 71-84, no diagnosis of neurological disorder), completed a non-enforced on-road driving assessment and detailed sensory-motor and cognitive testing. Self-reported and official data regarding crashes and traffic offences were collected for both the five years prior to the on-road assessment, and the 12 months following in order to determine whether those who received a Fail score on the on-road assessment had higher rates of negative driving events than those who passed (43 passed, 15 failed). No increase in adverse outcomes was found either retrospectively or prospectively for those who failed the on-road assessment. Similarly there were no significant differences in cognitive, sensory-motor, and demographic variables between those who passed and failed. Healthy older drivers who failed the on-road assessment did not show evidence of poorer driving behaviour even at the level of descriptive statistics.

INTRODUCTION

Older drivers are a rapidly growing section of the population at increased risk of being injured and killed in car accidents (McKnight & McKnight, 1999; OECD, 2001). Health concerns linked to unsafe driving in older drivers include chronic physical conditions (Dobbs, Caprio Triscott, & McCraken, 2004; Marottoli, Cooney, Wagner, Doucette, & Tinetti, 1994), visual attention and processing deficits (Ball & Owsley, 1991, 1996), and age-related cognitive decline and dementia (Cooper, Tallman, Tuokko, & Beattie, 1993; Dobbs, Heller, & Schopflocher, 1998; Johansson et al., 1996; McKnight & McKnight, 1999). As the percentage of older drivers in the population increases it is becoming increasingly important to have reliable and valid assessment tools to assess driving ability in this group.

Driving assessors regularly use both off- and on-road measures to assess driving ability (Korner-Bitensky, Bitensky, Sofer, Man-Son-Hing, & Gelinas, 2006). Off-road measures range from tests of visual acuity, neuropsychological tests, computerized cognitive test batteries, and complex driving simulators. On-road driving assessments are generally used as the "gold standard" measure of driving safety and often serve as the dependent variable in studies attempting to find predictors of driving safety. However, it has also been suggested that on-road assessments act as proxies for more critical measures, such as adverse driving outcomes in real-world driving situations (Owsley, Ball, Sloane, Roenker, & Bruni, 1991).

Many countries and states impose some degree of age-based testing on older drivers. For example, California requires a vision and written knowledge test at age 70, while New Jersey has no age-based requirements for licence retention. In New Zealand a compulsory on-road assessment was required for drivers aged 80 and above prior to 2006. These older New Zealand drivers were 1.33 times more likely to be involved in an injury-causing accident in the following two years for every time that the test had to be re-sat in order to secure a pass (Keall & Frith, 2004). This evidence suggests that the on-road driving test outcome is related to actual on-road driving outcomes. After 2006 the on-road assessment was abolished and the responsibility for detecting driving problems in older drivers fell solely to general practitioners who are charged with providing a medical certificate at age 75, 80, and biennially thereafter for drivers to retain their licence. This prompted a study to determine what sensory-motor, cognitive and demographic variables were associated with real-world driving outcomes in a group of older drivers with no neurological impairment.

The current study is a 12-month follow-up of 58 drivers aged 70 years and over who were tested off-road on a range of sensory-motor and cognitive tests and on an on-road driving assessment with a Pass or Fail outcome. Participants were volunteers with no repercussions for receiving an on-road Fail score. This provided a unique opportunity to follow the driving behaviour of a group of drivers deemed *a priori* to be safe or unsafe. Firstly, the relationship between officially-reported retrospective traffic offences in the five years prior to the on-road assessment were compared to Pass and Fail on-road outcomes. Next, on-road Pass and Fail outcomes were compared to prospective officially- and self-reported crashes and traffic offences. The results of the initial off-road sensory-motor and cognitive testing were also compared to prospective crashes and traffic offences.

METHOD

Participants

Participants were 58 drivers (50% male) aged 71-84 years (mean = 77.5, SD = 4.2). They were a convenience sample recruited from a variety of community groups (e.g., churches, recreational groups), word of mouth, and advertisements placed in older-person's and health-oriented magazines distributed free in Christchurch, New Zealand. Exclusion criteria included a history of moderate to severe brain injury, neurological disease, and current psychiatric disorder. Inclusion criteria were chosen to reflect those older drivers who would likely be considered low-risk due to a lack of diagnosed neurological disease and would likely receive a medical fitness to drive certificate without referral for specialist driving assessment. Forty-three participants received a Pass on the on-road driving assessment and 15 received a Fail (25.9%). Both the initial and current studies were approved by the Upper South A Regional Ethics Committee.

Off-Road Assessment

Participants completed a number of cognitive tests including an IQ estimate (the Wechsler Test of Adult Reading, Wechsler, 2001), a standardized version of the Mini-Mental State Examination (Molloy & Standish, 1997), the Dementia Rating Scale-2 (Mattis, Jurica, & Leitten, 2001), Trail Making Test A and B, and a letter cancellation task. Visual acuity in each eye was also assessed using a Snellen chart. Participants also underwent assessment on a computerized battery of sensory-motor and cognitive tests (*SMCTests*) (Innes, Jones, Anderson, Hollobon, & Dalrymple-Alford, 2009; Jones & Donaldson, 1995; Jones, Donaldson, & Parkin, 1989) developed as a research tool for assessment of patients with neurological disorders and for assessment of driving ability (Innes et al., 2007). *SMCTests* assesses visuospatial function, visuoperception, reaction time, ballistic movement, visuomotor tracking, decision-making, visual search, sustained attention, divided attention, impulse control, and planning. Test stimuli were presented on a computer monitor and responses recorded using a face-valid system comprising a steering wheel, direction indicators, and a set of foot pedals (Figure 1).

Figure 1. *SMCTests* apparatus

On-Road Assessment

An on-road driving assessment was administered by a specialist occupational therapist and a driving instructor, both of whom were blind to performance on off-road testing. Participants drove the same 45-min route. Conditions included driving on public roads with traffic hazards such as single-lane roundabouts, dual-lane roundabouts, dual-lane roads, controlled and uncontrolled intersections, and changes in speed zone (i.e., 50km/h, 60km/h, and 80km/h sections). The assessors provided a consensus Pass or Fail score based on an 11-point Driving Scale (Innes et al., 2007).

Officially-Recorded Traffic Offences and Crashes

The New Zealand Transport Agency provided information regarding traffic offences for the period January 2002 through December 2008. This covers approximately 5 years prior to the on-road driving assessment until the completion of the 12 month follow-up period. Examples of the potential types of offences recorded are speeding, driving while intoxicated, reckless driving, and violating posted traffic signs. The New Zealand Ministry of Transport provided information for the same time period regarding police-reported injury crashes.

Follow-up Interview

Participants were interviewed by telephone 1 year (SD = 8.0 days) after their on-road driving assessment. They were asked to report the number of crashes and traffic offences committed during the 12 months following their assessment. Self-report was used to capture crashes not attended by police and traffic offences not reflected in Land Transport New Zealand data. For example, official data does not cover speed camera tickets. Speed camera tickets are issued by unattended speed radar detectors that photograph vehicle licence plates which are used to identify and post fines to registered owners of speeding vehicles.

Data Analysis

Fisher's Exact Test was used to determine whether the presence of an officially-reported traffic offence in the five years prior to the on-road assessment was different between subsequent on-road Pass or Fail outcome. The difference between Pass and Fail outcomes and officially- or self-reported traffic offence or crashes prospectively over 12 months was investigated in the same way. Data was then regrouped into those who had and had not had any reported crash or traffic offence in 12 months since their on-road driving assessment, and group membership was compared to scores on the initial sensory-motor and cognitive tests, and demographic data using t tests for normally distributed data and Mann-Whitney U tests for non-normally distributed data. The existence of at least one officially- *or* self-reported crash or offence was coded as a binary value for each participant.

RESULTS

At the first-12 month follow-up, 57 of 58 participants were still driving. One participant had ceased driving for several months due to a non-driving related injury but was intending to drive again when medically possible.

Officially-reported retrospective crashes and traffic offences

There were no officially-reported crashes in the five years prior to the on-road assessment. There were seven officially-reported traffic offences committed by six participants (one committed two offences) (Table 1). All seven traffic offences were for exceeding a posted speed limit of either 50 km/h or 100 km/h. All offences were committed by participants who went on to Pass the on-road assessment. There was no difference in the number of traffic offences reported between on-road Pass and Fail groups (Fisher's Exact Test, two-tailed $p = .15$).

Officially and self-reported traffic offences and crashes in subsequent 12 months

There were nine traffic offences across eight drivers in the 12-month period following the on-road assessment. Five offences were self-reported only (e.g., fixed camera speeding violations).

Three offences were both officially- and self-reported, and one was officially reported only. Two of the participants who committed an offence received an on-road Fail score (Table 1).

There were four self-reported crashes across four drivers, with none officially recorded. None of the people who reported crashes also had self- or officially-reported traffic offences. Only one of the accidents was committed by a driver who received an on-road Fail score.

There was no difference in the number of traffic offences or crashes between on-road Pass and Fail groups (Fisher's Exact Test, two-tailed $p = .63$).

Table 1. Number of participants with one or more retrospectively or prospectively reported crashes or offences (number in Pass and Fail on-road group)

	Type of traffic event	
	Participants with Crashes	Participants with Offences
Retrospective (5 yr)	0	6 (0 Fail, 6 Pass)
Prospective (12 mo)	4 (1 Fail, 3 Pass)	8 (2 Fail, 6 Pass)

Sensory-motor, cognitive, and demographic predictors of crashes and offences in subsequent 12 months

There were no significant differences found in sensory-motor and cognitive test performance and demographics between the 12 participants who experienced a crash or traffic offence in the 12 months following the off-road assessment and the remaining 46 who did not.

DISCUSSION

This is the first study to follow a group of older drivers *a priori* judged to be safe or unsafe based on a comprehensive on-road driving assessment. No relationship was found either between retrospective or prospective crashes and traffic offences and the on-road Pass or Fail score. Additionally, no relationship was found between cognitive and sensory-motor measures and prospective crashes and traffic offences. The lack of discernable differences in real-life driving between the 'safe' and 'unsafe' driver groups highlights difficulties in predicting driving safety. Research has shown that loss of a driver's licence in the elderly is associated with loneliness and immobility (Johnson, 1999). The low base-rate of traffic offences and crashes makes decisions regarding forfeiture of licences particularly fraught when the likely harm to the older driver versus harm to the community are taken into consideration.

Langford, Bohensky, Koppel, and Newstead (2008) recently compared driving fatalities of older drivers across two Australian states, one with compulsory medical and on-road assessments for drivers aged 80 and above, and one with no testing requirement. They found no difference in the number of fatalities of older drivers or other road users and concluded that age-based testing shows no demonstrable safety benefits. Another possibility for this result is that the on-road assessment used was not sensitive to the factors that predict real-world driving and those who had their licences forfeited were not more dangerous than those who did not. Another large sample study of older drivers found significant cognitive test predictors of prospective at-fault crashes over a period of five years (Ball et al., 2006). However, the driving ability of the participants in this study had not been assessed by the authors *a priori*. Allowing drivers judged to be unsafe using an on-road driving assessment to continue driving, as was the case in the present study, has to the best of our knowledge not been done in driving research literature. Thus the study provides a unique perspective on how dangerous those considered to be unsafe really are, beyond that provided by comparing populations with and without compulsory testing, or by relying solely on an on-road assessment to define safety. Had the results of the on-road assessments been disseminated to participants' doctors, 15 participants could have had their licences revoked. This equates to 15 participants who would have been left without personal transportation and the associated difficulties this imposes, and yet there were no differences in traffic incidents even at the level of descriptive statistics.

One weakness of the study is the relatively small number of participants and the low base-rate of crashes and offences. This weakens our ability to find significant differences between groups. Given the serious personal consequences for an older driver losing their licence, however, we should expect that those judged to be unsafe following a driving assessment should have noticeable increases in crashes and/or offences even at the level of descriptive statistics. Another weakness of the study is that the validity of the on-road assessment is unknown, despite similar assessments being routinely performed to assess driver safety throughout New Zealand and around the world. While a 24-month follow-up of this group which will provide additional data regarding driver safety is planned, the current 12-month data allows no discrimination between safe and unsafe groups.

ACKNOWLEDGEMENTS

The authors would like to acknowledge funding for this project from the Transport Research and Educational Trust Board, Accident Compensation Corporation, Christchurch Neurotechnology Research Programme, and University of Canterbury. This research was supported (in part) by a contract from the Health Research Council of New Zealand.

References

Ball, K. K., Roenker, D. L., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin, G., Raleigh, R., Joyce, J. J., Cissell, G. M., & Dube, T. (2006). Can high-risk older drives be identified through performance-based measures in a Department of Motor Vehicles setting? *Journal of the American Geriatrics Society, 54*, 77-84.

Cooper, P. J., Tallman, K., Tuokko, H., & Beattie, B. L. (1993). Vehicle crash involvement and cognitive deficit in older drivers. *Journal of Safety Research, 24*, 9-17.

Dobbs, A. R., Caprio Triscott, J. A., & McCraken, P. N. (2004). Considerations for assessment of medical competence to drive in older patients. *Geriatrics and Aging, 7*, 42-46.

Dobbs, A. R., Heller, R. B., & Schopflocher, D. (1998). A comparative approach to identify unsafe older drivers. *Accident Analysis & Prevention, 30*, 363-370.

Innes, C. R. H., Jones, R. D., Anderson, T. J., Hollobon, S. G., & Dalrymple-Alford, J. C. (2009). Performance in normal subjects on a novel battery of driving-related sensory-motor and cognitive tests. *Behavior Research Methods, 42*, 284-294.

Innes, C. R. H., Jones, R. D., Dalrymple-Alford, J. C., Hayes, S., Hollobon, S., Severinson, J., Smith, G., Nicholls, A., & Anderson, T. J. (2007). Sensory-motor and cognitive tests can predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences, 260*, 188-198.

Johansson, K., Bronge, L., Lungberg, C., Persson, A., Seiderman, M., & Vittanen, M. (1996). Can a physician recognize an older driver with increased crash risk potential? *Journal of the American Geriatrics Society, 44*, 1198-1204.

Johnson, J. E. (1999). Urban older adults and the forfeiture of a driver's license. *Journal of Gerontological Nursing, 25*, 12-18.

Jones, R. D., & Donaldson, I. M. (1995). Fractionation of visuoperceptual dysfunction in Parkinson's disease. *Journal of the Neurological Sciences, 131*, 43-50.

Jones, R. D., Donaldson, I. M., & Parkin, P. J. (1989). Impairment and recovery of ipsilateral sensory-motor function following unilateral cerebral infarction. *Brain, 112*.

Keall, M. D., & Frith, W. J. (2004). Association between older driver characteristics, on-road driving test performance, and crash liability. *Traffic Injury Prevention, 5*, 112-116.

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., & Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *American Journal of Occupational Therapy, 60*, 428-434.

Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008). Do age-based mandatory assessments reduce older drivers' risk to other road users? *Accident Analysis & Prevention, 40*, 1913-1918.

Marottoli, R. A., Cooney, L. M., Wagner, D. R., Doucette, J., & Tinetti, M. E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine, 121*, 842-846.

Mattis, S., Jurica, P. J., & Leitten, C. L. (2001). *Dementia Rating Scale - 2*. Lutz, Florida: Psychological Assessment Resources, Inc.

McKnight, A. J., & McKnight, A. S. (1999). Multivariate analysis of age-related driver ability and performance deficits. *Accident Analysis and Prevention, 31*, 445-454.

Molloy, D. W., & Standish, T. I. M. (1997). A guide to the Standardized Mini-Mental State Examination. *International Psychogeriatrics, 9* (Suppl. 1), 87-94.

OECD. (2001). *Ageing and transport: mobility and safety issues*. Paris: OECD Publications.

Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*, 403-415.

Wechsler, D. (2001). *Wechsler Test of Adult Reading*. San Antonio, Texas: Pearson Education, Inc.

11.4 Appendix D – Published Paper 4

ON-ROAD DRIVING ASSESSMENT ERRORS ASSOCIATED WITH PASS AND FAIL OUTCOMES FOR OLDER DRIVERS WITH COGNITIVE IMPAIRMENT

Petra Hoggarth[1,2], Carrie Innes[1,3], John Dalrymple-Alford[1,2,4], & Richard Jones[1,2,3,4,5]
[1] Van der Veer Institute for Parkinson's & Brain Research
[2] Department of Psychology, University of Canterbury
[3] Department of Medical Physics & Bioengineering, Christchurch Hospital
[4] Department of Medicine, University of Otago
[5] Department of Electrical & Computer Engineering, University of Canterbury
Christchurch, New Zealand
E-mail: petra.hoggarth@vanderveer.org.nz

Summary: Most on-road assessments do not make use of standardized scoring to determine driver safety. The current study sought to find a subset of driving errors that were related to on-road Pass and Fail outcomes in a group of 60 older drivers with cognitive impairment, and that were also considered important contributors to a Fail outcome by the driving specialist occupational therapist who administered the assessment. A number of useful errors were found that suggest that even a non-standardized driving assessment could incorporate a short list of driving errors that may assist in the determination of driving ability.

INTRODUCTION

On-road driving assessment is a commonly used tool to determine whether a driver with cognitive impairment or dementia is safe to continue driving. Some standardized assessments have been tested for inter-rater and test-retest reliability, with the former usually found to be moderate to high, and the latter in the moderate range (Hagge, 1994; Fitten et al., 1995; Romanowicz & Hagge, 1995; Hunt et al., 1997; Janke & Eberhard, 1998). Unfortunately, standardized on-road driving assessments are often not used as part of routine driving assessment. Korner-Bitensky et al. (2006) surveyed the driving assessment methods of 144 American and Canadian driving assessors. Ninety-four percent of respondents reported routinely using on-road assessments as part of their evaluation, yet only 24% used a standardized scoring system, and only 10% used a pre-defined cutoff score to define driving competency. Only two respondents reported using a standardized road test. The use of un-standardized on-road assessments in naturalistic research samples can make results difficult to generalize. Likewise, applying the results of studies that use standardized on-road assessments and scoring to real-world assessment practices can be equally as difficult.

The appeal of a non-standardized format is in the maximum flexibility it allows an assessor to draw on their often considerable clinical experience in deciding whether a person is a safe driver. For example, a standardized route cannot be used when an assessment is performed from a person's home and in their local area. A standardized on-road scoring system does not record all variables that may figure into an assessor's decision-making process, such as missed appointments that may indicate a memory problem. This flexibility in assessment may be particularly important when assessing people with cognitive impairment or dementia due to the idiosyncratic way impairments affect the abilities of individuals. We believe, however, that it is important that at least some standardization of on-road scoring is present in order that judgments made between drivers and between assessors are as equitable as possible.

For the current study, 60 drivers with suspected or diagnosed cognitive impairment or dementia were recruited. All completed an on-road driving assessment as used in New Zealand for assessing the impact of medical conditions on driving safety. This assessment did not use a standardized route or a predetermined rating of errors to determine a Pass or Fail outcome. Rather, the occupational therapist administering the assessment used a global decision-making process encompassing performance on the on-road assessment. The aim of the study was to find errors that occurred during the assessment that discriminated between those rated as on-road Pass and Fail. A detailed error list was constructed based on the results of a previous driving study. This was completed for each participant by both the driving instructor and occupational therapist who administered the assessment. Ideally, errors with a high degree of face validity to a driving assessor would be found that would be useful in discriminating between safe and unsafe drivers and could be used as part of the standard medical driving assessment. This would allow some form of objective scoring to be used in determining a Pass of Fail outcome for an assessment that currently does not have standardized scoring.

The study had three aims:

1. Determine the inter-rater agreement of commission of errors between the occupational therapist and driving instructor who administered the assessment.
2. Determine the errors most frequently cited by the occupational therapist as contributing to a Fail score.
3. Determine which errors discriminated between those who passed and failed the assessment and whether these errors were the same as those rated as most often contributing to a Fail score by the occupational therapist.

METHODS

Participants

Participants with a current full driver's licence were recruited from consecutive referrals to the Driving and Vehicle Assessment Service (DAVAS) at Burwood Hospital, Christchurch, New Zealand. Eligible referrals had suspected or diagnosed Alzheimer's dementia, memory or cognitive impairment listed as the reason for referral. These criteria were designed to include those participants who were likely to have Alzheimer's dementia or amnestic mild cognitive impairment.

Sixty participants were recruited: 36 males and 24 females with a mean age of 77.9 years (SD = 7.8, range 58–92). Participants gave informed consent. The study was approved by the Upper South A Regional Ethics Committee, Canterbury, New Zealand.

Driving Error List

A previous study investigated the on-road driving ability of 60 cognitively healthy older drivers (Hoggarth et al., 2010). An informal report including observed driving errors was compiled for each participant by the assessing occupational therapist. An error list with a standardized recording procedure was constructed based on the errors performed in this prior study. Noted errors were tallied and ranked, with the most common on-road errors forming the Driving Error List used for the current study. This list was given to the driving instructor of the prior study and two occupational therapists in order to add other errors that they commonly observed in people with dementia. Items in the final Driving Error List are presented in Table 1.

Testing Procedure

Participants completed an on-road driving assessment of around 45 minutes administered by an experienced driving occupational therapist and a driving instructor from the Driving and Vehicle Assessment Service at Burwood Hospital, Christchurch. Fifty-six drivers completed the on-road assessment in their own car, with the remaining four using an assessment centre car. The driving instructor sat in the passenger seat and provided directions for the driving route and maintained safety of the vehicle if required. The occupational therapist sat in the rear to assess driving behaviour. Twelve (20%) participants started the drive from their own homes, with the remainder 80% starting from the Hospital. Assessment from home was made for a variety of different reasons such as availability of a family member in order to drive the participant home if they received a Fail outcome.

The driving instructor and the occupational therapist completed the Driving Error List independently at the end of each assessment. Assessors ticked one box to indicate which errors were performed during the assessment and another box if the error was cited as a reason for a Fail rating for the on-road driving performance. The assessors could also record additional errors as they arose.As was normal procedure, a global decision of Pass or Fail was made by the occupational therapist based on the entirety of the participant's performance during the assessment.

Data Analysis

Intra-class correlations were computed using Cronbach's Alpha to assess reliability between the two raters. Errors identified by the occupational therapist as contributing to on-road Fail were summed and ranked by frequency. The relationship between each error and Pass and Fail groups was investigated using Fisher's Exact Test (for expected cell frequencies <5). The Phi Coefficient was computed as a measure of association of each error to the Pass and Fail rating. The Phi Coefficient is equivalent to a Pearson's r for a bivariate correlation.

RESULTS

Twenty-one of the 60 participants (35%) received a Fail score on the on-road assessment. There was a trend for the Fail group to have a higher mean age than the Pass group (mean pass = 76.4 years, mean fail = 80.8 years, $Z = -1.88$, $p = .06$). There was no difference in sex between groups (male = 1, female = 2: mean pass = 1.41, mean fail = 1.38, $Z = -.22$, $p = .83$), and no difference in Pass and Fail outcomes in those who started the on-road assessment from home versus the hospital ($Z = -.537$, $p = .59$). Those who failed the on-road performed more on-road errors than those who passed, (Fail mean = 11.1, Pass mean = 2.3, $Z = -6.40$, $p < .001$).

Eight additional unique errors not found on the error list were recorded: 'Approached intersections at excessive speed' (7 participants), 'Incorrect use of turning bays' (3 participants), 'Failure to stop at stop sign' (3 participants), 'Braking and accelerating at the same time' (3 participants), 'Excessive acceleration' (1 participant), 'Went to wrong side of the road when approaching intersections' (1 participant), 'Late moving into turning lane' (1 participant), and 'Went through red light' (1 participant). Since these errors were not present on the error list for systematic rating they were not included in the following calculations, but may be useful to be added to a revised error list.

Table 1 displays the frequency that each error was rated by the two raters and Cronbach's alpha.

Table 1. Cronbach's alpha and 95% confidence interval for agreement on commission of errors between raters

Type of Error	Frequency rated by occupational therapist	Frequency rated by driving instructor	Cronbach's alpha	95% confidence interval
Fails to follow pedestrian crossing rules	3	3	1.00	1.00 – 1.00
Decreased awareness of environment	21	20	0.98	.97 - .99
Decreased awareness of other road users	21	19	0.96	.94 - .98
Driving/starting in wrong gear	7	6	0.96	.93 - .97
Inappropriate gap selection	18	15	0.94	.98 - .96
Incorrect use of give way rules at intersections	14	13	0.92	.87 - .95
Driving below the speed limit	16	16	0.91	.84 - .94
Driving above the speed limit	29	27	0.89	.82 - .93
Incorrect indication at a roundabout	42	44	0.86	.77 - .92
Turned into incorrect lane on multi-lane road	4	7	0.83	.72 - .90
Inappropriate use of arrow traffic lights	2	1	0.80	.66 - .88
Driving too close to (or over) centre line	9	11	0.78	.64 - .87
Lack of scanning techniques	19	21	0.77	.62 - .86
Driving too close to (or over) left line	15	11	0.76	.60 - .86
Didn't react in time to situation/Incorrect action taken	13	8	0.76	.60 - .86
Lack of mirror use	21	24	0.76	.60 - .86
Problem cornering: speed or position	8	8	0.72	.54 - .84
Incorrect use of lanes in roundabout	3	4	0.71	.51 - .83
Fails to observe signs	8	9	0.69	.47 - .81
Immediate fail error (e.g. crash)	2	7	0.61	.34 - .77
Lack of blind spot check	17	23	0.59	.31 - .75
Following other cars too closely	12	3	0.56	.26 - .74
Didn't apply 12 second search	6	17	0.55	.24 - .73
Gear grinding/over-revving	5	1	0.50	.16 - .70
Stopping too closely behind cars	5	2	0.41	.01 - .65
Approaching intersections at excessive speed	2	6	0.36	-.07 - .62
Incorrect indication at an intersection	2	5	-0.11	-.86 - .34
Incorrect indication for lane changes	1	0	N/A	N/A
Fails to give way to pedestrians at intersection	1	0	N/A	N/A

There was a high degree of agreement between the raters with 11 error ratings with alphas of 0.80 or above. The error 'Incorrect indication at an intersection' had a negative alpha of -0.11 due to no agreement between the two ratings for any of the 7 participants who made this error.

The number of times an error was rated by the occupational therapist as contributing to a Fail score were ranked (Table 2). The two most frequently rated reasons for a Fail score were 'Decreased awareness of other road users' and 'Decreased awareness of environment', with this second reason being rated as a reason for all 21 participants who failed the assessment. The error 'Incorrect indication at a roundabout' had the highest occurrence in the sample as a whole (42 participants according to the occupational therapist's rating) but was only considered by the occupational therapist to have directly contributed to one person's Fail score.

Table 2. Frequency that errors were rated as contributing to a Fail outcome by the occupational therapist, and the statistical relationship of errors to Pass and Fail outcomes

Error	Frequency rated as contributing to the Fail score	Fisher's Exact Test *p* value	Phi coefficient
Decreased awareness of other road users	21	<.01	1.00
Decreased awareness of environment	20	<.01	0.93
Lack of scanning techniques	15	<.01	0.78
Inappropriate gap selection	13	<.01	0.74
Driving above the speed limit	13	<.01	0.48
Incorrect use of give way rules at intersections	12	<.01	0.67
Didn't react in time to situation / incorrect action taken	10	<.01	0.55
Driving below the speed limit	8	0.01	0.35
Driving too close to (or over) centre line	7	<.01	0.47
Driving too close to (or over) left line	6	<.01	0.54
Problem cornering: speed or position	6	<.01	0.53
Lack of mirror use	6	<.01	0.49
Lack of blind spot check	5	<.01	0.39
Following other cars too closely	5	0.09	0.24
Gear grinding/over-revving	3	<.01	0.41
Didn't apply 12 second search	3	0.02	0.34
Fails to observe signs	3	0.02	0.33
Incorrect use of lanes in roundabout	3	0.04	0.31
Stopping too closely behind cars	3	0.05	0.28
Driving/starting in wrong gear	2	0.23	0.17
Incorrect indication at an intersection	1	0.12	0.25
Incorrect indication at a roundabout	1	0.56	0.10
Fails to follow pedestrian crossing rules	1	1.00	-0.01
Immediate fail error (e.g. crash)	0	0.10	0.25
Incorrect indication for lane changes	0	0.35	0.18
Fails to give way to pedestrians at intersection	0	1.00	-0.10
Inappropriate use of arrow traffic lights	0	0.54	-0.14
Turned into incorrect lane on multi-lane road	0	0.29	-0.20

Eighteen errors were related to Pass and Fail outcomes (Table 2). All but four variables were in the expected direction of occurring more frequently in the Fail group. Three variables did not occur in the Fail group at all: 'Turned into incorrect lane on multi-lane road', 'Fails to give way to pedestrians at intersection', and 'Inappropriate use of arrow traffic lights'. The error 'Fails to follow pedestrian crossing rules' occurred once in the Fail group and twice in the Pass group.

The errors most related to a Fail outcome are shown by the largest phi coefficients and are similar to the occupational therapist's self-rated contribution of error to Fail scores (Table 2). Of the top 10 errors rated as contributing to Fail scores by the occupational therapist, 8 were ranked in the top ten by phi coefficient for errors associated with Fail scores: 'Decreased awareness of other road users', 'Decreased awareness of environment', 'Lack of scanning techniques', 'Inappropriate gap selection', 'Incorrect use of give way rules at intersections', 'Didn't react in time to situation / incorrect action taken', 'Driving too close to (or over) left line', and 'Driving above the speed limit'.

DISCUSSION

Inter-rater agreement on the commission of driving errors was high for many items. This indicates that many of the listed errors were identified independently by each rater. Discrepancies in some ratings could be due to the threshold of each rater in defining when an error occurred, and also due to the driving instructor's dual task of guiding the drive as well as maintaining safety of the vehicle, with the occupational therapist able to concentrate solely on driving behaviour.

The two measures most commonly cited as contributing to a Fail outcome were not errors per se, but rather a judgment made by the occupational therapist that a driver exhibited decreased awareness of the environment and/or other road users. The fact that these two awareness measures were rated for all but one Fail outcome suggests that a single measure of decreased awareness could be used instead of two separate ones. This result shows the importance of maintaining items which are not simply based on the performance of an error but rather on an interpretation based on a pattern of errors that indicate decreased awareness.

There was a high level of agreement between the subjective ratings of the occupational therapist as to which errors were especially important to a Fail outcome, and the degree to which the frequency of these errors (whether or not they were rated as important) were associated with on-road Fail scores. This suggests that a select choice of errors may be acceptable to assessors who prefer a flexible assessment procedure.

There are limitations to this study. Firstly, the occupational therapist may have been biased to recall more driving errors for drivers whom she had decided would receive Fail outcomes, thus the number of errors recalled may be inflated in the Fail group. The number of errors selected for the error list was limited and additional errors would likely have been useful, particularly errors more related to the types of difficulties that might be expected to be more common in drivers with cognitive impairment – confusion for example.

Finding ways to bridge the gap between researchers who ideally prefer a standardized on-road assessment and on-road assessors who frequently prefer a more flexible and less standardized approach could be achieved, in part, by including a checklist of errors that have been found to relate to on-road driving outcome in older adults with cognitive impairment. This study found a number of potentially useful error measures that were related to on-road Pass and Fail outcomes and were also rated as useful contributors to identifying people with on-road fail outcomes by an occupational therapist. Including an evidence-based error list as part of a non-standardized driving assessment would allow for systematic collection of error information that may contribute to more reliable Pass and Fail outcomes.

ACKNOWLEDGEMENTS

The authors would like to acknowledge funding for this project from the Canterbury Medical Research Foundation, Accident Compensation Corporation, Christchurch Neurotechnology Research Programme, the University of Canterbury, the Maurice and Phyllis Paykel Trust, and Universities New Zealand - Te Pōkai Tara.

REFERENCES

Fitten, L. J., Perryman, K. M., Wilkinson, C. J., Little, R. J., Burns, M. M., Pachana, N., Mervis, R., Malmgren, R., Siembieda, D. W., & Ganzell, S. (1995). Alzheimer and vascular dementias and driving: a prospective road and laboratory study. *Journal of the American Medical Association, 273*, 1360-1365.

Hagge, R. A. (1994). The California driver performance evaluation project: an evaluation of a new driver licensing road test. Sacramento, CA: California Department of Motor Vehicles.

Hoggarth, P. A., Innes, C. R. H., Dalrymple-Alford, J. C., Severinsen, J. E., & Jones, R. D. (2010). Comparison of a linear and a non-linear model for using sensory-motor, cognitive, personality, and demographic data to predict driving ability in healthy older adults *Accident Analysis and Prevention, 42*, 1759-1768.

Hunt, L. A., Murphy, C. F., Carr, D., Duchek, J. M., Buckles, V., & Morris, J. C. (1997). Reliability of the Washington University Road Test. *Archives of Neurology, 54*, 707-712.

Janke, M. K., & Eberhard, J. W. (1998). Assessing medically impaired older drivers in a licensing agency setting. *Accident Analysis and Prevention, 30*, 347-361.

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., & Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *American Journal of Occupational Therapy, 60*, 428-434.

Romanowicz, P. A., & Hagge, R. A. (1995). An evaluation of the validity of California's driving performance evaluation road test. Sacramento, CA: California Department of Motor Vehicles.

11.5 Appendix E – Published Paper 5

PROCEEDINGS of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design

SELF-RATED DISTRESS RELATED TO MEDICAL CONDITIONS IS ASSOCIATED WITH FUTURE CRASHES OR TRAFFIC OFFENCES IN OLDER DRIVERS

Petra Hoggarth[1,2], Carrie Innes[1,3], John Dalrymple-Alford[1,2,4], & Richard Jones[1,2,3,4,5]
[1] Van der Veer Institute for Parkinson's & Brain Research
[2] Department of Psychology, University of Canterbury
[3] Department of Medical Physics & Bioengineering, Christchurch Hospital
[4] Department of Medicine, University of Otago
[5] Department of Electrical and Computer Engineering, University of Canterbury
Christchurch, New Zealand
E-mail: petra.hoggarth@vanderveer.org.nz

Summary: Ageing is associated with the development of medical conditions, both acute and chronic. The aim of this study was to determine whether medical factors were associated with subsequent self- and officially-reported crashes and traffic offences in a group of cognitively healthy older drivers. We surveyed medical conditions, medications taken for these conditions, and the amount of subjective distress associated with medical conditions in a group of 56 drivers aged 72-85 years for a period of 24 months. We also compared exposure to driving at baseline to the number of crashes or offences at 24 months. We found no relationship between the number of medical conditions or medications taken and whether a participant had a crash or offence. However, those who reported more subjective distress associated with their condition/s were more likely to have a crash or offence during the study period. Drivers who had a crash or offence also had a higher mean driving exposure. However, there was no relationship between reported distress and driving exposure which indicates that these may be independent risk factors for experiencing a crash or traffic offence.

INTRODUCTION

Dementia, cerebrovascular disease, heart disease, and high blood pressure have all been linked with poorer driving safety in older people (McGwin et al., 2000; Sims et al., 2000; Brown & Ott, 2004; Anstey et al., 2005; Sagberg, 2006; Ott et al., 2008). Comorbidity of chronic conditions increases the number of self-reported crashes in older drivers over 12-months (Marottoli et al., 1994). Molnar et al. (2007) report that older drivers who endorsed being "bothered a great deal by diabetes mellitus" were more likely to have had crashes. This suggests that subjective feelings of distress associated with medical conditions in general may influence driving safety.

Serious at-fault car crashes are a useful measure of driving safety, but are a rare event. A New Zealand study found an incidence rate of 0.8% for police-reported serious crashes (fault not determined) in a sample of over 39,300 adults aged 80 and over during a two year period (Keall & Frith, 2004a). Self-reported crashes, however, have a higher base-rate at around 4-8% per year (Marottoli et al., 1994; Sullman & Baas, 2004), and may be a useful outcome measure for indicating which drivers are performing unsafe driving behaviours.

Traffic offences have a higher base-rate than crashes. For the 12 months ended 30 June 2008, New Zealand police issued 760,720 speeding offences to a licensed population of just over 3 million drivers (New Zealand Police, 2008; Ministry of Transport, 2009). This represents a speeding offence for 24% of drivers, although this is an over-estimation as the data do not account for multiple offences for a single individual. Importantly, committing traffic offences has been associated with an increased likelihood of having crashes, even in older drivers (Rajalin, 1994; Parker et al., 1995; Cooper, 1997; Keall & Frith, 2004a). People observed driving at higher speeds on public roads were significantly more likely to have a state-recorded history of crashes and traffic offences over the previous seven years (Wasielewski, 1984). Therefore, a record of traffic offences may help in identifying drivers who are at increased risk for crashes.

In addition to the main aim of this study in exploring the relationship of medical factors and crashes and offences, we were also interested in the effect of driving exposure. Higher driving exposure has been linked with a greater likelihood of experiencing a crash (Owsley et al., 1998; Ball et al., 2006). Conversely, other studies have found that those older drivers driving fewer km were more likely to have a crash (Janke, 1991; Hakamies-Blomqvist et al., 2002; Keall & Frith, 2004b; Langford et al., 2006; Alvarez & Fierro, 2008).

In the current study, we followed 56 drivers aged 70 and older for 24 months with annual interviews collecting data about driving behavior and the presence of medical and psychological conditions. The drivers had no diagnosis of cognitive disorder at baseline. The aims of the study were to determine if crashes and traffic offences in a group of older drivers were related to:

- The number of medical conditions endorsed.
- The number of medical conditions endorsed for which a medication was taken.
- The amount of subjective distress due to medical conditions.
- The amount of driving exposure.

METHODS

Participants

Participants were recruited for a previous study investigating off-road predictors of on-road driving ability (Hoggarth et al., 2010). A convenience sample of 60 current drivers was recruited from churches, recreational groups, word of mouth, and advertisements placed in two free local health magazines in Christchurch. Participants were aged 70 to 84 years with 10 men and 10 women in each of three age groups (70-74, 75-79, and 80+ years). Exclusion criteria included a history of moderate to severe brain injury, diagnosed neurological or cognitive disorder (i.e. stroke or dementia), severe musculoskeletal disease, and current psychiatric disorder. Participants undertook a 3-hour off-road testing session that included a computerized sensory-motor and cognitive test battery (*SMCTests*™), personality measures, and standardized cognitive tests (Hoggarth et al., 2010). Participants completed a formal but unenforced 45-minute on-road driving assessment where the performance of each was rated as Pass or Fail (on-road assessors were blind to the results of all off-road testing). Participants continued to drive regardless of the outcome of the on-road assessment.

Participants from this original study were invited to take part in a 24-month follow-up study (the current study) using an annual 30-minute telephone interview. Officially-recorded traffic offence data from the New Zealand Transport Agency and police-reported crash data from the Ministry of Transport was collected. Participants gave informed consent, and both studies were approved by the Upper South A Regional Ethics Committee, Canterbury, New Zealand. One participant refused the follow-up study and one refused to give access to official crash and offence records. By 24 months, one participant had died and one could not be located, leaving interview data and full crash/offence data for 24 months for 56 participants. By the end of 24 months the mean age of the sample was 78.7 years (SD = 4.1) with a range of 72 to 85 years (27 males and 29 females). All remaining 56 participants were still active drivers at the 24 month interview.

Medical Condition Data

At each annual interview, participants were asked whether they had been diagnosed or treated for various medical conditions, psychiatric conditions, eye conditions, or possible indicators of declining health in the past 12 months (see Table 1). If a participant endorsed a condition they were asked whether they currently took a medication for this condition (rated as 'yes' or 'no') and subjective distress was assessed by asking how 'bothered' they were by the condition on a daily basis: not at all, a little, or a great deal.

Driving Exposure Data

As part of a previous study, participants logged their odometer readings prior to, and following, each driving trip taken during a seven day period. Additional information about longer trips taken in the last year was used to estimate driving exposure over the previous 12 months.

Crash and Traffic Offence Data

Crash data provided by the Ministry of Transport comprised crashes to which police were called. Traffic offence data provided by the New Zealand Transport Agency comprised offences issued in person by police officers. Parking offences were not included in this data. At 12- and 24-month interviews participants self-reported involvement in crashes and receipt of traffic tickets in the preceding 12 months. These self-reported data were collected to detect adverse events including crashes that were not attended by police, and traffic offences not collected on the official record such as roadside speed radar offences.

Data Analysis

One or more instances of either a crash or an offence reported either by a participant or an official source by the end of 24-month follow-up period formed the 'crash or offence' versus 'no-crash or offence' binary dependent variable. The number of medical conditions endorsed and the number of medical conditions requiring medication were summed for each participant. The amounts that participants reported being 'bothered' by each medical condition were assigned an ordinal rating of 0 for 'not at all', 1 for 'a little', 2 for 'a great deal' and were summed.

Mann-Whitney *U* tests were used to determine whether the number of medical conditions, the number of medications taken, the amount that participants were bothered by their medical conditions and the amount of driving performed were related to crash or offence involvement. Detailed examination of whether individual medical conditions were related to crashes and offences was not performed due to insufficient data across participants.

RESULTS

In total, 16 of 56 participants had self- or officially-reported crashes or offences over the 24-month period (10 participants with an offence only, 5 participants with a crash only, one participant with a crash and an offence). There were no officially-reported crashes for the 24-month follow-up. All officially-reported offences were also self-reported and all were for exceeding the speed limit. Descriptive statistics for medical conditions are provided in Table 1.

Table 1. Self-reported medical conditions medical conditions and indicators of declining health at the 12- and 24-month interviews for 56 initially cognitively healthy participants

	12 month follow-up			24 month follow-up		
Medical Condition	People reporting illness	People taking medication	People bothered by their condition (those rated > 0)	People reporting illness	People taking medication	People bothered by their condition (those rated > 0)
Arthritis	30	12	12	32	10	13
High blood pressure	31	31	0	27	27	0
High cholesterol	22	18	0	22	18	0
Cataracts	18	0	2	22	0	5
Heart disease	16	14	2	18	18	1
Surgery	10	0	1	11	2	2
Cancer	11	2	2	8	2	0
Osteoporosis	9	8	1	10	8	1
Fall	5	2	3	6	1	3
Diabetes	6	5	0	6	5	0
Glaucoma	4	4	0	4	4	1
Thyroid problems	2	2	0	3	3	0
Anxiety	2	1	0	3	1	1
Macular degeneration	2	1	0	2	1	1
Stroke	2	2	0	1	1	1
Depression	1	0	0	6	4	2
Broken bones	1	0	0	4	1	2
Dementia	1	0	0	1	0	0
Sleep apnea	1	1	1	1	1	0
Head injury	0	0	0	4	0	0
Parkinson's	0	0	0	1	1	1
Multiple sclerosis	0	0	0	0	0	0
Diabetic retinopathy	0	0	0	0	0	0
Retinal detachment	0	0	0	0	0	0

The most common medical conditions at 24 months were arthritis (57%), high blood pressure (48%), high cholesterol (39%), cataracts (39%), and heart disease (32%), with similar rates at 12 months. The single condition that distressed participants the most was arthritis, with 40% of those with arthritis at 12 months and 41% at 24 month reporting being bothered either a little or a great deal. There were also notable increases in the incidence of depression, from one person at 12 months to six at 24 months. Four people had broken bones at 24 months compared to one at 12 months; and head injuries increased from zero at 12 months to four at 24 months. These changes could indicate declining physical or emotional health of the sample over time.

Between those with and without crashes or offences, there was no difference in the number of medical conditions endorsed at 24 months ($z = -0.09$, $p = .93$), and no difference in the number of medications taken at 24 months ($z = -0.67$, $p = .50$). At 24 months, participants who had a crash or offence reported significantly more distress related to their medical condition/s than participants who had not had a crash or offence ($z = -2.01$, $p = .04$). Participants who had a crash or offence drove a higher annual mean number of km at baseline (crash or offence mean = 18,661 km, no crash or offence mean = 7,893 km, $z = -2.20$, $p = .03$). There was no significant correlation between annual km driven and distress associated with medical conditions ($r_s = -0.02$, $p = .86$).

DISCUSSION

No association was found between the number of medical conditions endorsed or the number of medications taken and whether a participant had experienced a crash or offence during the previous 24 months. However, the relationship between reports of subjective distress related to medical conditions at 24 months and the incidence of crashes or offences suggests that subjective physical or emotional discomfort associated with medical illness may affect a person's likelihood of an adverse driving event. Increased distress could be related to increased seriousness of the condition that could negatively impact on driving behaviour. Increased distress could also relate to increased distraction or reduced mobility caused by the condition, for example pain or restriction of movement caused by arthritis.

Molnar et al. (2007) found that older drivers who reported being bothered a great deal my diabetes were more likely to have had previous crashes. The current study supports the impact of subjective distress on adverse driving outcomes in a group of generally cognitively healthy older drivers. Options for amelioration of distressing symptoms, such as pain relief for those with arthritis, could have a positive impact on an older person's ability to drive safely.

The occurrence of a crash or offence over the following 24-month period was also related to a higher annual number of km driven measured at the baseline testing session. The difference was substantial with the crash or offence group on average driving over twice as much as the no crash or offence group. The fact that distress associated with medical conditions and driving exposure were not linearly related indicates that the two variables may be independent risk factors for crashes or offences

There are limitations in this study. Firstly, when initially recruited the participants were a sample of older drivers who would be considered representative only of a general older age sample that lacks diagnosed cognitive impairment. By the end of the 24-month period, one participant had been diagnosed with dementia, one with Parkinson's disease, and two had had a stroke. Thus, the sample began to resemble the general older population more closely as time progressed.

Secondly, it is likely that specific types of medical conditions or medications play more important roles in driving safety than merely the tallied number of conditions or medications, but the sample size was too small to provide enough power for a detailed investigation of this. Finally, the conflation of crashes and traffic offences is not ideal but it could be argued that both sit on a continuum of driving behaviours we would wish to minimize if possible. Future investigations into the relationship between at-fault crashes and traffic offences in larger samples would be valuable to determine if the two can be combined, particularly if a system for weighting the importance of crashes and offences could be devised.

With longevity showing signs of increase in many societies, it is likely that higher percentages of future cohorts of older drivers will be living, and driving, with the effects of chronic disease. Subjective ratings of distress related to these conditions may be a useful way to determine which older drivers are more likely to experience an adverse driving event. It may also be important to recognize that predominantly healthy drivers who drive greater distances are more likely to be involved in a crash or traffic offence.

ACKNOWLEDGEMENTS

The authors acknowledge funding for this project from the Transport Research and Educational Trust Board, Accident Compensation Corporation, Christchurch Neurotechnology Research Programme, and the University of Canterbury, the Maurice and Phyllis Paykel Trust, and Universities New Zealand - Te Pōkai Tara.

REFERENCES

Alvarez, F. J., & Fierro, I. (2008). Older drivers, medical condition, medical impairment and crash risk. *Accident Analysis and Prevention, 40*, 55-60.

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review, 25*, 45-65.

Ball, K. K., Roenker, D. L., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin, G., Raleigh, R., Joyce, J. J., Cissell, G. M., & Dube, T. (2006). Can high-risk older drives be identified through performance-based measures in a Department of Motor Vehicles setting? *Journal of the American Geriatrics Society, 54*, 77-84.

Brown, L. B., & Ott, B. R. (2004). Driving and dementia: a review of the literature. *Journal of Geriatric Psychiatry and Neurology, 17*, 232-240.

Cooper, P. J. (1997). The relationship between speeding behaviour (as measure by violation convictions) and crash involvement. *Journal of Safety Research, 28*, 83-95.

Hakamies-Blomqvist, L., Raitanen, T., & O'Neill, D. (2002). Driver ageing does not cause higher accident rates per km. *Transportation Research Part F, 5*, 271-274.

Hoggarth, P. A., Innes, C. R. H., Dalrymple-Alford, J. C., Severinsen, J. E., & Jones, R. D. (2010). Comparison of a linear and a non-linear model for using sensory-motor, cognitive, personality, and demographic data to predict driving ability in healthy older adults *Accident Analysis and Prevention, 42*, 1759-1768.

Janke, M. K. (1991). Accidents, mileage, and the exaggeration of risk. *Accident Analysis and Prevention, 23*, 183-188.

Keall, M. D., & Frith, W. J. (2004a). Association between older driver characteristics, on-road driving test performance, and crash liability. *Traffic Injury Prevention, 5*, 112-116.

Keall, M. D., & Frith, W. J. (2004b). Older driver crash rates in relation to type and quantity of travel. *Traffic Injury Prevention, 5*, 26-36.

Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk - a replication of low mileage bias. *Accident Analysis and Prevention, 38*, 574-578.

Marottoli, R. A., Cooney, L. M., Wagner, D. R., Doucette, J., & Tinetti, M. E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine, 121*, 842-846.

McGwin, G. J., Sims, R. V., Pulley, L., & Roseman, J. M. (2000). Relations among chronic medical conditions, medications, and automobile crashes in the elderly: a population-based case-control study. *American Journal of Epidemiology, 152*, 424-431.

Ministry of Transport. (2009). *Motor Vehicle Crashes in New Zealand 2008*. Wellington: Ministry of Transport.

Molnar, F. J., Marshall, S. C., Man-Son-Hing, M., Wilson, K. G., Byszewski, A. M., & Stiell, I. (2007). Acceptability and concurrent validity of measures to predict older driver involvement in motor vehicle crashes: an Emergency Department pilot case-control study. *Accident Analysis and Prevention, 39*, 1056-1063.

New Zealand Police. (2008). *Annual Report for the year ended 30 June 2008*. Wellington.

Ott, B. R., Heindel, W. C., Papandonatos, G. D., Festa, E. K., Davis, J. D., Daniello, L. A., & Morris, J. C. (2008). A longitudinal study of drivers with Alzheimer disease. *Neurology, 70*, 1171-1178.

Owsley, C., Ball, K. K., McGwin, G., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association, 279*, 1083-1088.

Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38*, 1036-1048.

Rajalin, S. (1994). The connection between risky driving and involvement in fatal accidents. *Accident Analysis and Prevention, 26*, 555-562.

Sagberg, F. (2006). Driver health and crash involvement: a case-control study. *Accident Analysis and Prevention, 38*, 28-34.

Sims, R. V., McGwin, G., Allman, R. M., Ball, K. K., & Owsley, C. (2000). Exploratory study of incident vehicle crashes among older drivers. *Journal of Gerontology, 55A*, M22-M27.

Sullman, M. J. M., & Baas, P. H. (2004). Mobile phone use amongst New Zealand drivers. *Transportation Research Part F, 7*, 95-105.

Wasielewski, P. (1984). Speed as a measure of driver risk: observed speeds versus driver and vehicle characteristics. *Accident Analysis and Prevention, 16*, 89-103.

11.6 Appendix F – The 14-Item Driving Anger Scale

Driving experiences scale

Date:…………………

Instructions: Imagine that each situation described below was actually happening to you and rate the amount of anger that would be provoked.

none at all	a little	some	much	very much
1	**2**	**3**	**4**	**5**

1.	Someone is weaving in and out of traffic.	☐
2.	A slow vehicle on a mountain road will not pull over and let people by.	☐
3.	Someone backs right out in front of you without looking.	☐
4.	Someone runs a red light or stop sign.	☐
5.	You pass a radar speed trap.	☐
6.	Someone speeds up when your try to pass him/her.	☐
7.	Someone is slow in parking and is holding up traffic.	☐
8.	You are stuck in a traffic jam.	☐
9.	Someone makes an obscene gesture toward you about your driving.	☐
10.	Someone honks at you about your driving.	☐
11.	A bicyclist is riding in the middle of the lane and is slowing traffic.	☐
12.	A police officer pulls you over.	☐
13.	A truck kicks up sand or gravel on the car you are driving.	☐
14.	You are driving behind a large truck and you cannot see around it.	☐

Please check: Did you write a number at the end of each statement?

11.7 Appendix G – Driving Behaviour Questionnaire

Driver Behaviour Questionnaire (DBQ)

Instructions: Following is a list of common mistakes and violations that people make when driving. For each statement you are asked to indicate how often, if at all, these things have happened to you say over the last year. You are to answer each statement with one of six responses: never, hardly ever, occasionally, quite often, frequently, nearly all the time. Your answers do not have to be precise, merely your best estimate and you are not obliged to answer all of the questions if you would prefer not to.

0 = Never, 1 = Hardly ever, 2 = Occasionally, 3 = Quite often, 4 = Frequently, 5 = Nearly all the time

Realize you have no recollection of the road along which you have just been traveling
0 –1 – 2 – 3 – 4 - 5

Brake too quickly on a slippery road, or steer the wrong way into a skid
0 –1 – 2 – 3 – 4 - 5

Become impatient with a slow driver and pass them when a solid yellow line means you are in a no passing lane
0 –1 – 2 – 3 – 4 - 5

Switch on one thing, such as the headlights, when you meant to switch on something else, such as the wipers
0 –1 – 2 – 3 – 4 - 5

Do not notice for a few moments that the traffic light has turned green
0 –1 – 2 – 3 – 4 - 5

Drive especially close to the car in front as a signal to its driver to go faster or get out of the way
0 –1 – 2 – 3 – 4 - 5

Intending to drive to destination A, you suddenly notice that you are on the road to destination B, perhaps because B is your more usual destination
0 –1 – 2 – 3 – 4 - 5

Attempt to overtake someone you had not noticed to be signaling a right turn
0 –1 – 2 – 3 – 4 - 5

Cross an intersection knowing the traffic lights have already turned against you
0 –1 – 2 – 3 – 4 - 5

Have an aversion to a particular class of road user, and indicate your hostility by whatever means you can

0 –1 – 2 – 3 – 4 - 5

Miss give way signs and narrowly avoid colliding with traffic having right of way
0 –1 – 2 – 3 – 4 - 5

Forget where you left your car in a car park
0 –1 – 2 – 3 – 4 - 5

Underestimate the speed of an oncoming vehicle when overtaking
0 –1 – 2 – 3 – 4 - 5

Queuing to turn left onto main road, you pay such close attention to the main stream of traffic that you nearly hit the car in front
0 –1 – 2 – 3 – 4 - 5

Fail to check rearview mirror before pulling out, changing lanes, etc.
0 –1 – 2 – 3 – 4 - 5

Drive even though you realize that you may be over the legal blood-alcohol limit
0 –1 – 2 – 3 – 4 - 5

Get into the wrong lane approaching a roundabout or intersection
0 –1 – 2 – 3 – 4 - 5

On turning left, nearly hit a cyclist who has come up on your inside
0 –1 – 2 – 3 – 4 - 5

Fail to notice pedestrians crossing when turning into a side road
0 –1 – 2 – 3 – 4 - 5

Angered by another driver's behaviour, you give chase with the intention of giving him or her a piece of your mind
0 –1 – 2 – 3 – 4 - 5

Hit something when reversing that you had not previously seen
0 –1 – 2 – 3 – 4 - 5

Get involved in unofficial 'races' with other drivers
0 –1 – 2 – 3 – 4 - 5

Misread signs and take the wrong turning off a roundabout
0 –1 – 2 – 3 – 4 - 5

Disregard the speed limits late at night or early on the morning
0 –1 – 2 – 3 – 4 - 5

11.8 Appendix H – Driving Habits Questionnaire Used for the Healthy Driver Follow-up Study

Modified Driving Habits Questionnaire (DHQ)

Dead
Month/year:
Cause:

Current Driving
1. Do you currently drive?
1 = yes *(go to #A)*   0 = no *(go to #2, #3, #A only)*

2. Why did you stop driving?

3. When is the last time you drove? ____________________ (month/year)
(If within 1 year, go to question #25)

A. What forms of transport are you currently using?
☐ Drive in own car
☐ Driven as passenger by friend/family member
☐ Taxi
☐ Motor scooter / motorcycle
☐ Bicycle
☐ Bus
☐ Walking
☐ Other____

B. In your opinion are you now driving the same amount, more, or less than one year ago?
Same
More
Less

C. Have you taken any driving lessons in the last year?
Yes
No

D. Has your GP asked you about your driving in the last year?
Yes
No

E. Have you been diagnosed with or treated for any of the following conditions over the past 12 months?

Medical Condition	Meds	Bothered by		
		Not at all	A little	A great deal
Heart disease	Y / N			
Cancer	Y / N			
Stroke	Y / N			
Parkinson's	Y / N			
Dementia	Y / N			
High blood pressure	Y / N			
High cholesterol	Y / N			
Thyroid problems	Y / N			
Sleep apnoea	Y / N			
Diabetes	Y / N			
Multiple sclerosis	Y / N			
Broken bones	Y / N			
Arthiritis	Y / N			
Depression	Y / N			
Anxiety	Y / N			
Osteoporosis	Y / N			
Surgery	Y / N			
Fall	Y / N			
Head injury	Y / N			
Eye Conditions:				
Cataracts	Y / N			
Macular degeneration	Y / N			
Diabetic retinopathy	Y / N			
Glaucoma	Y / N			
Retinal detachment	Y / N			

4. Do you wear glasses or contact lenses when you drive?
1 = yes 0 = no

5. Do you wear a seatbelt when you drive?
1 = always 2 = sometimes 3 = never

6. Which way do you prefer to get around?
3 = drive yourself 2 = have someone drive you 1 = use public transportation or a taxi

7. How fast do you usually drive compared to the general flow of traffic? Would you say:
5 = Much faster 4 = Somewhat faster 3 = About the same 2 = Somewhat slower
1 = Much slower

8. Has anyone suggested over the past year that you limit your driving or stop driving?
1 = yes 0 = no
9. How would you rate the quality of your driving? Would you say:
5 = Excellent 4 = Good 3 = Average 2 = Fair 1 = Poor

10. If you had to go somewhere and didn't want to drive yourself would you:
1 = Ask a friend or relative to drive you 2 = Call a taxi or take the bus
3 = Drive yourself regardless of how you feel 4 = Cancel or postpone your plans

5 = Other:

Avoidance

Now I'm going to ask some more specific questions about your driving habits.
17a) During the past 3 months, have you driven when it is raining?

_________Yes (**go to 17b**) __________No (**go to 17c**)

17b) Would you say that you drive when it is raining with:
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

17c) Do you deliberately avoid driving in the rain?

1________Yes _________No

18a) During the past 3 months, have you driven alone?

_________Yes (**go to 18b**) __________No (**go to 18c**)

18b) Would you say that you drive alone with: (Please check only **one** answer)
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

18c) Do you deliberately avoid driving alone?

1________Yes _________No

19a) During the past 3 months, have you parallel parked?

_________Yes (**go to 19b**) __________No (**go to 19c**)

19b) Would you say that you parallel park with: (Please check only **one** answer)
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

19c) Do you deliberately avoid parallel parking?

1________Yes _________No

20a) During the past 3 months, have you made right-hand turns across oncoming traffic?

_________Yes (**go to 20b**) __________No (**go to 20c**)

20b) Would you say that you make right- handed turns in traffic with:
(Please check only **one** answer) across oncoming traffic?
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

20c) Do you deliberately avoid making right-hand turns?

1________Yes _________No

21a) During the past 3 months, have you driven on motorways or highways?

_________Yes (**go to 21b**) __________No (**go to 21c**)

21b) Would you say that you drive on motorways or highways with: (Please check only **one** answer)
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

21c) Do you deliberately avoid driving on motorways or highways?

1________Yes _________No

22a) During the past 3 months, have you driven on high-traffic roads?

_________Yes (**go to 22b**) __________No (**go to 22c**)

22b) Would you say that you drive on
high-traffic roads with: (Please check only **one** answer)
5 = No difficulty at all 4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

22c) Do you deliberately avoid driving on high traffic roads?

1________Yes _________No

23a) During the past 3 months, have you driven in rush-hour traffic?

_________Yes (**go to 23b**) __________No (**go to 23c**)

23b) Would you say that you drive
in rush hour traffic with: (Please check only **one** answer)
5 = No difficulty at all   4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

23c) Do you deliberately avoid driving in rush-hour traffic?

1________Yes _________No

24a) During the past 3 months, have you driven at night?

_________Yes (**go to 24b**) __________No (**go to 24c**)

24b) Would you say that you drive at night with: (Please check only **one** answer)
5 = No difficulty at all   4 = A little difficulty 3 = Moderate difficulty 2 = Extreme difficulty

24c) Do you deliberately avoid driving at night?

1________Yes _________No

Crashes and Citations

25. How many accidents have you been involved in over the past 12 months when you were the driver? Please tell me the number of all accidents, whether or not you were at fault.

____ accidents

26. How many accidents have you been involved in over the past 12 months when you were the driver where the police were called to the scene?

____accidents

27. How many times in the past 12 months have you been pulled over by the police, regardless of whether you received a ticket?

____times

28. How many times in the past 12 months have you received a traffic ticket (other than a parking ticket) , regardless of whether or not you think you were at fault?

____times

Driving Space
29. During the past year, have you driven in your immediate neighbourhood?
1 = yes   0 = no

30. During the past year, have you driven to places beyond your neighbourhood?
1 = yes   0 = no

31. During the past year, have you driven to neighbouring towns?
1 = yes   0 = no

32. During the past year, have you driven to more distant towns?
1 = yes 0 = no

33. During the past year, have you driven to places outside of the South island?
1 = yes 0 = no

11.9 Appendix I – Road Code Questions Used in the Dementia and Driving Study

Driving questions

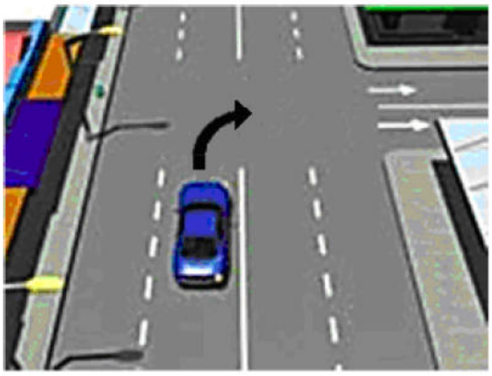
1) When turning right at a roundabout, you must:

A. not indicate at any time on the roundabout.
B. indicate left before entering the roundabout.
C. indicate right before entering the roundabout.
D. indicate right when leaving the roundabout.

---

2) When turning right from a two-laned road into a one-way street that has two lanes, you must turn into the:

A. right-hand lane.
B. left-hand lane.
C. lane that has the least vehicles.
D. lane that gives you the most direct route.



---

3) Before making a lane change to your left, you must signal for at least 3 seconds and:

A. use your vehicle brake lights to warn other road users.
B. use your vehicle hazard lights to warn other road users.
C. check your blind spot before moving left.
D. ensure that there are no oncoming vehicles.

4) When coming up to a pedestrian crossing, what MUST you do?

A. Slow down and be ready to stop for pedestrians.
B. Speed up before pedestrians cross.
C. Sound the horn on your vehicle to warn pedestrians.
D. Slow down to 30 km/h.

---

5) When should you use the 2-second rule?

A. When reversing out of a driveway.
B. When following other vehicles at the same speed.
C. When giving way at a pedestrian crossing.
D. When using the indicators on your vehicle.

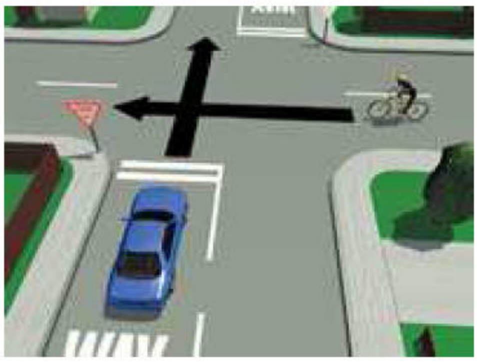
---

6) If you have to drive at a slow speed, which may hold up other vehicles, what should you do?

A. Keep as close as you can to the centre of the road.
B. Keep as close as you can to the left side of the road.
C. Make the traffic behind you slow down to the speed that you are driving at.
D. Drive down the middle of the road so that any vehicles behind you can pass on the left.

---

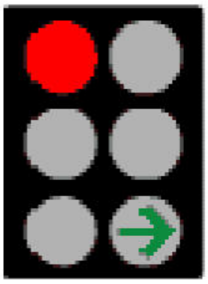
7) You are the driver of the blue car. Do you:



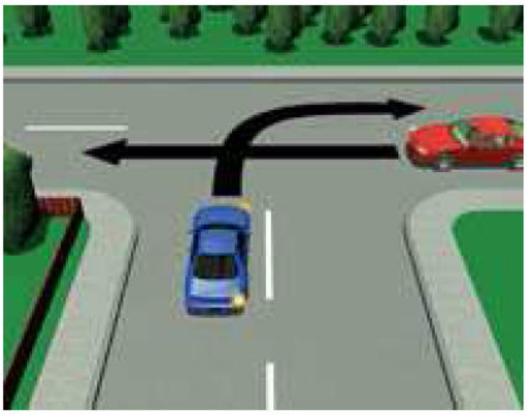
A. keep driving as the cyclist has to stop?
B. stop and give way to the cyclist?
C. keep driving as the cyclist has to give way to larger vehicles?
D. sound the warning device to warn the cyclist?

8) What may you do at traffic signals if there is a green arrow pointing to the right and a red light showing at the same time?

A. You may turn right.
B. You may go straight ahead.
C. You must stop until all lights turn green.
D. You may turn left or right.

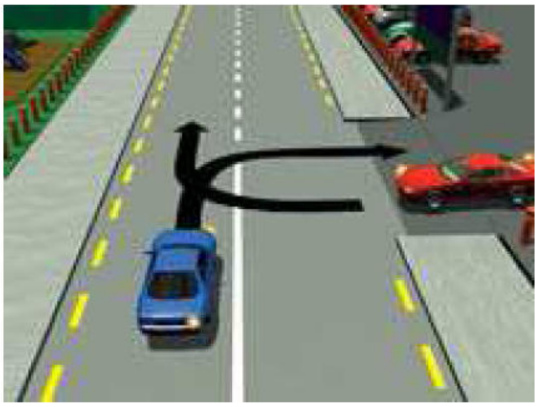


Does the driver of the blue car have to give way?
9)



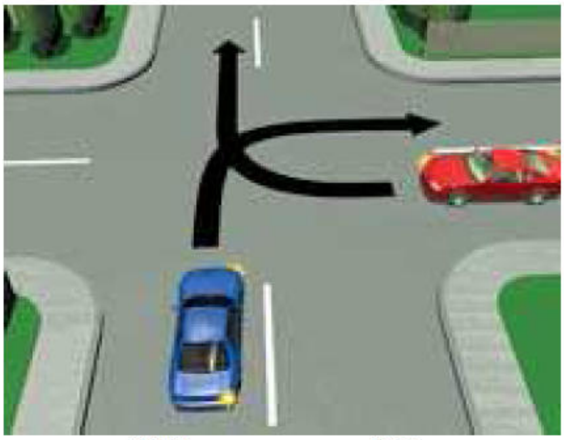
YES **NO**

Does the driver of the blue car have to give way?
10)



YES **NO**

Does the driver of the blue car have to give way?
11)



YES **NO**

Does the driver of the blue car have to give way?
12)

YES **NO**

---

13) The 'speed limit' means the:

A. slowest speed you can drive in good conditions.
B. average speed you can drive in good conditions.
C. fastest speed you can drive in good conditions.
D. fastest speed you can drive, except when overtaking.

14) The number on this sign shows:

A. the fastest speed to drive around the curve.
B. the speed in the right-hand lane.
C. the safest speed to drive around the curve.
D. the distance to the next intersection.

---

15) What do these road markings mean?

A. Turning vehicles should enter turning lane at arrows and avoid striped area.
B. Turning vehicles should avoid all road markings and use the left-hand lane.
C. All vehicles can make a U-turn at these road markings.
D. Turning vehicles drive straight over all road markings and wait to turn right.

11.10 Appendix J – Driving Scale Used During On-Road Assessments

Table 9-1. The driving scale used during to assign scores following driving assessment in the Healthy Older Drivers and Dementia and Driving studies

Driving Score	Outcome	Label
0	Fail	No ability
1	Fail	Basic skills only
2	Fail	Extremely inferior
3	Fail	Very poor
4	Fail	Poor
5	Fail	Borderline
6	Pass	Fair
7	Pass	Satisfactory
8	Pass	Good
9	Pass	Very good
10	Pass	Flawless

11.11 Appendix K – Form Used for Listing of Driving Errors Completed for Participants in the Dementia and Driving Study

Driving Errors for on-road assessment

Referral name and number: Date:

Errors	Error present ✓	Persisted despite instruction ✓	Contributed to fail judgment ✓
Incorrect indication at a roundabout			
Incorrect indication at an intersection			
Driving below the speed limit			
Driving above the speed limit			
Incorrect use of give way rules (controlled or uncontrolled intersections)			
Inappropriate gap selection			
Driving/starting in wrong gear			
Gear grinding/over-revving/wrong gear selected			
Lack of blind spot check			
Following other cars too closely			
Stopping too closely behind cars			
Driving too close to (or over) centre line			
Driving too close to left line			
Turned into incorrect lane on multi-lane road			
Incorrect use of lanes in roundabout			
Incorrect indication for lane changes (short duration, no indication)			
Fails to give way to pedestrians at intersection			
Fails to follow pedestrian crossing rules			
Inappropriate use of arrow lights on traffic signals			
Decreased awareness of environment			
Decreased awareness of other road users			
Lack of mirror use			
Didn't apply 12 second search			
Lack of scanning techniques			
Didn't reacts in time to situation, incorrect action taken			
Fails to observe signs			
Problem cornering: speed / position (specify)			
Immediate fail error (e.g. crash) (specify)			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			
Other (specific):			

Insight Yes Limited No

Awareness of driving problems Yes Limited No

Receptiveness to feedback Yes Limited No

Likely to benefit from lessons Yes No

Pass or Fail: P F

Driving Score: 0 1 2 3 4 5 6 7 8 9 10

11.12 Appendix L – Information Sheet for the Healthy Older Drivers Study

RESEARCH STUDY INFORMATION

‘Computerized driving tests for predicting driving in healthy older adults’

Computerized driving tests for predicting driving in healthy older adults

What is this research study about?

You are invited to take part in helping us with our research. This is an information sheet about a study which aims to assess the accuracy and predictive value of our off-road driving assessment system. The study will determine how well off-road tests can predict on-road driving ability. We are currently looking for a group of healthy older adults whom we can assess in order to determine the normal range of performance expected for an older age group. The off-road tests will assess your ability to sustain attention on a task, plan a sequence of events, follow a target, and measure your reaction times. It is also hoped that the tests will help identify specific problems underlying an inability to drive safely for individual people. This is especially important now that GPs are responsible for predicting whether older adults are safe to drive. We want to develop a system that is fair and less stressful for older drivers who may be referred for a driving assessment, and that gives the same results that we would get from an on-road test.

The study will take part in two sessions. The first session will be around 2 - 2 ½ hours long and will take place at the Van der Veer Institute for Parkinson's and Brain Research located at 66 Stewart St close to Christchurch Hospital. This session will include a brief interview (5-10 minutes), some standard tests of mood, memory and reading (approximately 60 minutes), and then an assessment using a computerized driving system called the Canterbury Driving Assessment Tool (*CanDAT*) (approximately 60 minutes). The *CanDAT* has a steering wheel, indicators, and pedals. You do not need to know how to use a computer to undertake the assessment. You will need to look at some images on a computer screen and to read a vision chart as a distance of 3 meters – so remember to bring any glasses that you will need with you.

The second session will be around 1 hour and will comprise an on-road driving assessment. The on-road driving assessment will commence at Burwood Hospital and will be undertaken by an occupational therapist and a driving instructor. You may use your own car for this assessment, as long as it has room in the back seat for the occupational therapist to sit. The assessment will commence in quiet streets and then move through a number of different driving situations which will include single-lane roundabouts, dual-lane roundabouts, dual-lane roads, controlled intersections (give-way, stop sign, and traffic light controlled) uncontrolled intersections, and changes in speed zone.

We will try to schedule the two sessions of testing with you as close together as we can. At the completion of the two assessments you will be reimbursed to the value of $50 (by posted cheque) to cover any petrol/vehicle expenses for your visits to the Van der Veer Institute and Burwood Hospital.

Participation in this study will not affect the status of your driving license. The occupational therapist will discuss the results of the on-road assessment with you and if she believes it is necessary, will recommend that you attend some driving lessons to focus on areas where your driving may need improvement. The occupational therapist will only make such suggestions if she believes there are aspects of your driving that could be addressed to make your driving safer. **Your GP will not be informed of your performance in the driving assessment.**

In the unlikely event that some of your scores indicate the possibility of health issues such as depression, anxiety, or severe cognitive problems, information may, with your prior consent, be passed on to your GP.

Your participation in this study is entirely voluntary. If you do agree to take part you are free to withdraw at any time from the study, without having to give a reason and this will not affect any future care or treatment.

A possible extension to this study is to follow participants over a period of 1 and 2 years to answer questions about the previous year's driving (such as continued driving or accident occurrence) either over the phone or via a postal questionnaire. The commencement of this study will be subject to ethical approval through the Upper South Ethics Committee before any participant is contacted.

Who is running this research study?

This study has been developed by researchers at the Van der Veer Institute for Parkinson's and Brain Research. The Principle Investigator (Petra Hoggarth) is completing this research as part of the requirements for a Master of Arts degree for the Department of Psychology, University of Canterbury. The study has received ethical approval from the Upper South A Regional Ethics Committee. We aim to recruit 60 older drivers in the age groups of 70-74, 75-79, and 80 years and over (20 participants for each age group). The study is expected to start in June 2007 and testing will continue until around January 2008.

All records will be kept confidential during and after the study, and you will be identified in these records only by an assigned subject code number. The information gathered will only be used for the purposes of the study. Only the researchers of this project will have access to records associated with the study, which will be kept in safe storage at the Van der Veer Institute for up to 10 years. No material which could personally identify you will be used in any reports on this study. Although individual results will not be provided, if you wish, a summary of our findings will be sent to you on completion of this study. Please note that there will be a delay between the completion of data collection and publication of the results.

The results of this study will help to refine the ability of the computerized tests to predict driving ability, and the tests may be commercialised for sale through the Canterbury District Health Board.

What do you need to do?

We would greatly value your help. You are welcome to discuss the study with your friends/family or the research staff before making a decision on whether to take part. If you are interested in taking part, please phone or email the Principle Investigator whose details are at the bottom of the next page. If you do agree to participate, please complete the questionnaires that are included with this information sheet. Please bring all the completed questionnaires with you when you attend the first assessment session.

If you have any questions about the study you can contact Petra Hoggarth, either by phone at the Van der Veer Institute (Ph 378 6095 – please leave a message if I am not in) or by email (petra.hoggarth@vanderveer.org.nz). Thank you for considering this request.

Yours sincerely,

Petra Hoggarth BA(Hons)	**Richard Jones PhD**
Master of Arts and Clinical Psychology student	Research Associate Professor

Van der Veer Institute for
Parkinson's & Brain Research

66 Stewart Street	Phone:	03 378 6095
Christchurch	Email:	petra.hoggarth@vanderveer.org.nz

11.13 Appendix M – Consent Form for the Healthy Older Drivers Study

VAN DER VEER INSTITUTE for Parkinson's & Brain Research

UC UNIVERSITY OF CANTERBURY Te Whare Wānanga o Waitaha CHRISTCHURCH NEW ZEALAND

Consent Form

'Computerized driving tests for predicting driving in healthy older adults'

I have read and understood the information sheet dated 19/06/2007 for volunteers taking part in the study designed to improve the prediction of driving ability in healthy older drivers. I have had the opportunity to discuss this study. I am satisfied with the answers I have been given.

I have had the opportunity to use whanau support or a friend to help me ask questions and understand the study.

I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time and this will in no way affect my future healthcare.

I understand that my participation in this study is confidential and that no material which could identify me will be used in any reports on this study.

I have had time to consider whether to take part.

I know whom to contact if I have any questions about the study.

I consent to being contacted in the future and completing a questionnaire. YES / NO

I wish to receive a copy of the results of this study YES / NO

I give consent for information of serious health issues that may arise during testing (such as the possible existence of depression, anxiety, or severe cognitive problems) to be passed on to my GP YES / NO

I ____ hereby consent to take part in this study.

(print full name)

Date:______________________________

Signature:____

Researchers: Petra Hoggarth BA(Hons), Carrie Innes PhD, Richard Jones PhD, John Dalrymple-Alford PhD, and Julie Severinsen BHSc(OT)

Phone: 378 6095

Project explained by: Petra Hoggarth

Project role: Principle Investigator

Signature:

Date:

Ethnicity

Which ethnic group do you belong to? *Mark the space or spaces which apply to you.*	Ko tēhea momo tāngata e whai pānga atu ana koe? *Tohua te katoa o raro nei e hāngai ana ki a koe.*
◯ NZ European	◯ Pākehā
◯ Māori	◯ Māori
◯ Samoan	◯ Hāmoa
◯ Cook Island Maori	◯ Māori Kuki Airani
◯ Tongan	◯ Tonga
◯ Niuean	◯ Niue
◯ Chinese	◯ Hainamana
◯ Indian	◯ Īnia
◯ other (such as *DUTCH, JAPANESE, TOKELAUAN). Please state:*	◯ tētahi atu (pērā i *TATIMANA, HAPANĪHI, TOKELAU). Tuhia mai:*

11.14 Appendix N – Shapiro-Wilk *W* Test Scores for Independent Variables in the Healthy Older Drivers Study

Shapiro-Wilk *W* test scores are presented for demographic measures, cognitive measures, and *SMCTests* measures in Table 9-2 and Table 9-3.

Table 9-2. Shapiro-Wilk *W* test score for demographic and cognitive measures for the Healthy Older Drivers study

Test Measure	W Statistic	*p*-value (****p***<.05)
Gender	0.637	**.000***
Age (years)	0.936	**.004***
Age Grouping (70-74, 75-79, 80+)	0.793	**.000***
Handedness (left, right)	0.271	**.000***
Years of education	0.960	**.047***
Occupation code	0.822	**.000***
Years of driving	0.967	.099
Kms driven last 12 months	0.394	**.000***
Vision		
Left eye	0.506	**.000***
Right eye	0.761	**.000***
Binocular	0.825	**.000***
Road sign test (no correct)	0.751	**.000***
Mini-Mental State Exam	0.870	**.000***
Geriatric Depression Scale	0.864	**.000***
Beck Anxiety Inventory	0.842	**.000***
Driving Anger Scale time 1	0.979	.370
Driving Anger Scale time 2	0.959	**.040***
Big Five Inventory		
Extraversion	0.992	.970
Agreeableness	0.973	.198
Conscientiousness	0.962	.057
Neuroticism	0.967	.109
Openness to experience	0.985	.668
Trail Making Test A (s)	0.923	**.001***
Trail Making Test B (s)	0.818	**.000***
Wechsler Test of Adult Reading	0.934	**.003***
Dementia Rating Scale -2 AEMSS[1]	0.979	.399

[1]Age and education adjusted mean scale score

Table 9-3. Shapiro-Wilk *W* test score for *SMCTests* measures for the Healthy Older Drivers study –

Test Measure	W Statistic	*p*-value (***p*<.05)**
Footbrake and Clutch Test		
Mean reaction time (ms)	0.911	**.000***
Mean movement time (ms)	0.951	**.017***
Total reaction and movement times (ms)	0.941	**.006***
Ballistic Movement Test		
Reaction time, right hand (ms)	0.827	**.000***
Reaction time, left hand (ms)	0.861	**.000***
Reaction time, grand mean (ms)	0.957	**.034***
Movement time, right hand (ms)	0.779	**.000***
Movement time, left hand (ms)	0.901	**.000***
Movement time, grand mean (ms)	0.899	**.000***
Total reaction and movement times, right hand (ms)	0.816	**.000***
Total reaction and movement times, left hand (ms)	0.925	**.001***
Total reaction and movement times, grand mean (ms)	0.938	**.004***
Peak velocity, right hand (ms)	0.975	.255
Peak velocity, left hand (ms)	0.981	.484
Peak velocity, grand mean (ms)	0.981	.456
Tracking Tests		
Sine tracking run 1 error (mm)	0.891	**.000***
Sine tracking run 2 error (mm)	0.745	**.000***
Random tracking run 1 error (mm)	0.829	**.000***
Random tracking run 2 error (mm)	0.881	**.000***
Arrows Perception Test		
Number of arrows correct	0.614	**.000***
Divided Attention Test		
Tracking error (mm)	0.854	**.000***
Number of arrows correct	0.693	**.000***
Omission of arrows response	0.110	**.000***
Visual Search Test		
Mean reaction time (ms)	0.987	.779
Number correct	0.969	.131

Continued on following page

		Continued from previous page
Test Measure	W Statistic	*p*-value (***p*<.05**)
Complex Attention test		
Reaction time (ms)	0.943	**.007***
Movement time (ms)	0.947	**.011***
Total mean movement and reaction times (ms)	0.972	.190
Reaction time standard deviation (ms)	0.832	**.000***
Movement time standard deviation (ms)	0.542	**.000***
Number of lapses errors	0.227	**.000***
Number of invalid trials	0.362	**.000***
Planning Test		
Lateral road position error (mm)	0.975	.247
Duration of positional faults (s)	0.917	**.001***
Distance travelled (m)	0.772	**.000***
Intersection safety margin (mm)	0.971	.166
Number of hazards hit	0.926	**.001***
Number of crashes	0.811	**.000***

11.15 Appendix O – Measures Excluded from Offering to the Healthy Older Drivers BLR Model

Table 9-4. Measures that were excluded on pragmatic grounds and reasons for exclusion for the Healthy Older Drivers study

Measures not offered to BLR model	Reason for not offering
Occupation code	Would not be used to predict in public health setting
Years of driving	Would not be used to predict in public health setting
Years of education	Would not be used to predict in public health setting
Km driven last 12 months	Would not be used to predict in public health setting
Handedness	Would not be used to predict in public health setting
Driving Anger Scale time 1 and 2	Explorative only, and would not be used to predict in public health setting
Big Five Inventory - Extraversion	Explorative only, and would not be used to predict in public health setting
Big Five Inventory - Agreeableness	Explorative only, and would not be used to predict in public health setting
Big Five Inventory - Conscientiousness	Explorative only, and would not be used to predict in public health setting
Big Five Inventory - Neuroticism	Explorative only, and would not be used to predict in public health setting
Big Five Inventory - Openness to experience	Explorative only, and would not be used to predict in public health setting
Vision – left eye	Binocular vision measure used instead
Vision – right eye	Binocular vision measure used instead
Footbrake and Clutch, Mean movement time	Effect in the opposite from expected direction (Fails moved faster)
Ballistic Movement, Reaction time, right hand	Grand mean measure used instead of individual hand scores, also effect in opposite from expected direction (Fails had a faster reaction time)
Ballistic Movement, Reaction time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Total time, right hand	Grand mean measure used instead of individual hand scores, also effect in opposite from expected direction (Fails had a faster total time)
Ballistic Movement, Total time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, left hand	Grand mean measure used instead of individual hand scores
Complex Attention, Number of invalid trials	Effect in the opposite from expected direction (Fails had few invalid trials)
Complex Attention, Movement time standard deviation	Effect in the opposite from expected direction (Fails had a lower standard deviation)
Divided Attention, Omission of arrows response	Effect in the opposite from expected direction (Fails omitted fewer responses)
Arrows perception, Non-response	Too little variation in the data to make it a sensitive predictor
Planning, Duration of positional faults	Effect in the opposite from expected direction (Fails maintained positional faults for less time)
Planning, Number of hazards hit	Effect in the opposite from expected direction (Fails ran into fewer hazards)

11.16 Appendix P – Information Sheet for the Dementia and Driving Study

RESEARCH STUDY INFORMATION

‘Factors associated with driving in people with memory problems**’**

Factors associated with driving in people with memory problems

What is this research study about?

You are invited to take part in helping us with our research. This is an information sheet about a study which aims to determine how well tests performed off-road can predict on-road driving ability. We are currently looking for a group of adults who have been referred to Burwood Hospital for a driving assessment due to concerns about memory problems.
There are two ways to be involved in this study (Study A and Study B). They are discussed separately below.

Study A

Involvement in Study A will involve taking part in one or two sessions of testing in addition to the standard driving assessment. The tests will assess your memory, ability to plan and perform tasks, and other areas of brain functioning. It is also hoped that the tests will help identify specific problems underlying driving ability. This is especially important because our current methods for assessing driving for people with memory problems are not very accurate. We want to develop a system that is fair and less stressful for drivers who may be referred for a driving assessment, and that gives similar results to that we would get from an on-road driving assessment.

What does the research study involve?

The study takes part in two sessions or around 2 to 2.5 hours each that happen on separate days. Sessions can take place either in your own home or at the Van der Veer Institute for Parkinson's and Brain Research. The sessions will include a brief interview, as well as tests for various brain functions. The tests will all involve sitting at a table and answering questions, performing activities with a pencil and paper, or using objects such as blocks.
At the completion of the testing session you will be reimbursed with a $30 petrol voucher to pay for any travel expenses or for your time associated with attending the session.
For the remainder of the study you need to attend your driving assessment session at Burwood Hospital as normal. It is important that the extra testing session happens before your driving assessment at Burwood Hospital.
Your participation in this study is entirely voluntary. If you do agree to take part you are free to withdraw at any time from the study, without having to give a reason and this will not affect any future care or treatment. You are more than welcome to bring a support person with you to the testing session.

What information do we collect?

We will collect information from you such as your age and other basic details. We will also collect information from the tests about how your brain is working. The researcher Petra Hoggarth will also need to talk to someone who knows you well, such as a spouse, child, friend or neighbour with whom you have regular contact. This person will be asked some questions about how you are functioning in daily life. This can be done in a telephone conversation or in person.
The researcher Petra Hoggarth would also receive information regarding your driving assessment at Burwood Hospital.

In order to get a good record of driving behaviour we would like to access information about police reported accidents and injuries held by the Ministry of Transport, and information regarding traffic violations such as speeding tickets from Land Transport New Zealand for the period 1 June 2003 to the 30th September 2010. This covers a period of around five years before your on-road driving assessment. If you do not want us to access these records we will not do so, and instead will use information that you tell us yourself.
With your permission, data from this study may be used in future related studies, which have been given ethical approval from a Health & Disability Ethics Committee.

Study B

Involvement in Study B requires no extra testing other than attending your scheduled driving assessment as normal. Study B does not provide as much useful research about driving, but participation in Study B will still help us understand the relationship between off- and on-road driving assessments. By agreeing to be in Study B researchers will have access to the results of your off- and on-road driving assessment results from Burwood Hospital. We will also have access to other basic information taken at the assessment such as your age and results of a brief vision test. You will be asked to sign a consent form when you attend your driving assessment if you are willing to take part in Study B.

With your permission, data from this study may be used in future related studies, which have been given ethical approval from a Health & Disability Ethics Committee.

Who is running this research study?

The study has been developed by researchers at the Van der Veer Institute for Parkinson's and Brain Research. The Primary Researcher (Petra Hoggarth) is completing this research as part of the requirements for a PhD degree for the Department of Psychology, University of Canterbury. The study has received ethical approval from the Upper South A Regional Ethics Committee. We aim to recruit 60 drivers. The study is expected to start in July 2008 and testing may continue until around September 2010.
All records will be kept confidential during and after the study, and you will be identified in these records only by an assigned subject code number. The information gathered will only be used for the purposes of the study. Only the researchers of this project will have access to records associated with the study, which will be kept in safe storage at the Van der Veer Institute for up to 10 years. No material which could personally identify you will be used in any reports on this study. Although individual results will not be provided, if you wish, a summary of our findings will be sent to you on completion of this study. Please note that there will be a delay between the completion of data collection and publication of the results. The results of this study will help to refine the predictive ability of computerized tests used at Burwood Hospital for driving assessment, and the tests may be commercialised for sale through the Canterbury District Health Board.

Your rights and an ACC statement

If you have questions or concerns about your rights as a participant in this research study you can contact an independent health and disability advocate:
This is a free service provided under the Health and Disability Commissioner Act.

Telephone (NZ wide): 0800 555 050
Free Fax (NZ wide): 0800 2787 7678 (0800 2 SUPPORT)
Email (NZ wide): advocacy@hdc.org.nz

In the unlikely event of a physical injury as a result of your participation in this study, you may be covered by ACC under the Injury, Rehabilitation and Compensation Act. ACC cover is not automatic and your case will need to be assessed by ACC according to the provisions of the 2002 Injury Prevention Rehabilitation and Compensation Act. If your claim is accepted by ACC, you still might not get any compensation. This depends on a number of factors such as whether you are an earner or non-earner. ACC usually provides only partial reimbursement of costs and expenses and there may be no lump sum compensation payable. There is no cover for mental injury unless it is a result of physical injury. If you have ACC cover, generally this will affect your right to sue the investigators. If you have any questions about ACC please contact your nearest ACC office or feel free to ask the researcher for more information before you take part in this study.

What do you need to do?

We would greatly value your help. You are encouraged to discuss the study with your friends/family or the research staff before making a decision on whether to take part. If you are interested in taking part, please phone or email the Primary Researcher whose details are at the bottom of the page.
If you agree to be in Study A, It is important that the extra testing sessions for this study occur before your visit to Burwood Hospital for driving assessment.
If you have any questions about either Study A or B you can contact Petra Hoggarth, either by phone at the Van der Veer Institute (Ph 378 6095 – please leave a message if I am not in) or by email (petra.hoggarth@vanderveer.org.nz). Thank you for considering this request.

Yours sincerely,

Petra Hoggarth BA(Hons)
PhD and Clinical Psychology student

Richard Jones PhD
Research Associate Professor

Van der Veer Institute for
Parkinson's & Brain Research
66 Stewart Street
Christchurch

Phone: 03 378 6095
Email: petra.hoggarth@vanderveer.org.nz

11.17 Appendix Q – Shapiro-Wilk W Test Scores for the n=60 Dementia and Driving Study

Table 9-5. Shapiro-Wilk *W* test score for demographic and *SMCTests* measures for the n=60 Dementia and Driving study

Test Measure	W Statistic	*p*-value (***p*<..05*)**
Gender	0.622	**.000***
Age (years)	.958	**.036***
Handedness (left, right)	.374	**.000***
Ballistic Movement Test		
Reaction time, right hand (ms)	.785	**.000***
Reaction time, left hand (ms)	.720	**.000***
Reaction time, grand mean (ms)	.776	**.000***
Movement time, right hand (ms)	.919	**.001***
Movement time, left hand (ms)	.825	**.000***
Movement time, grand mean (ms)	.890	**.000***
Total reaction and movement times, right hand (ms)	.845	**.000***
Total reaction and movement times, left hand (ms)	.797	**.000***
Total reaction and movement times, grand mean (ms)	.817	**.000***
Peak velocity, right hand (ms)	.967	.098
Peak velocity, left hand (ms)	.972	.179
Peak velocity, grand mean (ms)	.969	.135
Tracking Tests		
Sine tracking run 1 error (mm)	.894	**.000***
Sine tracking run 2 error (mm)	.743	**.000***
Random tracking run 1 error (mm)	.878	**.000***
Random tracking run 2 error (mm)	.870	**.000***
Arrows Perception Test		
Arrows Perception Test	.663	**.000***
Omission of arrows response	.370	**.000***
Divided Attention Test		
Tracking error (mm)	.930	**.002***
Number of arrows correct	.748	**.000***
Omission of arrows response	.631	**.000***

Continued on following page

Continued from previous page

Test Measure	W Statistic	*p*-value (******p***<..05**)
Complex Attention test		
Reaction time (ms)	.463	**.000***
Movement time (ms)	.948	**.013***
Total mean movement and reaction times (ms)	.970	.146
Reaction time standard deviation (ms)	.970	.141
Movement time standard deviation (ms)	.854	**.000***
Number of lapses errors	.564	**.000***
Number of invalid trials	.573	**.000***
Planning Test		
Lateral road position error (mm)	.739	**.000***
Duration of positional faults (s)	.774	**.000***
Distance travelled (m)	.893	**.000***
Intersection safety margin (mm)	.867	**.000***
Number of hazards hit	.893	**.000***
Number of crashes	.831	**.000***

11.18 Appendix R – Independent Measures Excluded from Offering to the n=60 Dementia and Driving Model with Reasons for Exclusion

Table 9-6. Measures that were excluded on pragmatic grounds and reasons for exclusion for the n=60 Dementia and Driving study

Measures not offered to BLR model	Reason for not offering
Handedness	Would not be used to predict in public health setting
Ballistic Movement, Reaction time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Reaction time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Total time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Total time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, left hand	Grand mean measure used instead of individual hand scores
Planning, Distance travelled	Effect in the opposite from expected direction (Fails drove further during the test)
Arrows perception, Non-response	No effect in either direction – results the same between Pass and Fail groups

11.19 Appendix S – Shapiro-Wilk W Test Scores for Independent Variables for n=32 Dementia and Driving Study

Shapiro-Wilk *W* test scores are presented for demographic measures, driving measures, cognitive measures, and *SMCTests* measures in Table 9-7, Table 9-8, and Table 9-9.

Table 9-7. Shapiro-Wilk *W* test scores for demographic and driving measures for n=32 Dementia and Driving Study

Test Measure	W Statistic	*p*-value (***p*<..05*)*
Gender	.585	**.000***
Age (years)	.976	.668
Handedness (left, right)	.391	**.000***
Diagnosis (1 = MCI, 2 = Alzheimer's)	.540	**.000***
Years of Education	.910	**.011***
Years of Driving	.905	**.008***
Road code questions	.952	.163
Demerit points earned in the previous 5 years	.490	**.000***
Forms of transport used regularly:		
drive own car	.334	**.000***
taken as passenger	.265	**.000***
taxi	.334	**.000***
bicycle	.172	**.000***
bus	.602	**.000***
walking	.565	**.000***
Wear glasses when driving	.565	**.000***
Self-rated quality of driving (ordinal)	.821	**.000***
People who during the last year have driven:		
to neighbouring towns	.478	**.000***
to distant towns	.615	**.000***
to places outside of the South Island	.172	**.000***

Table 9-8. Shapiro-Wilk *W* test scores for cognitive measures for n=32 Dementia and Driving Study

Test Measure	W Statistic	*p*-value (****p*<..05*)
Geriatric Depression Scale	.826	**.000***
Wechsler Test of Adult Reading	.977	.698
Mini-Mental State Exam	.911	**.012***
ADAS-Cog Total score	.972	.557
VOSP Incomplete Letters (z-score)	.777	**.000***
VOSP Silhouettes (z-score)	.966	.389
Colour-Word Interference, Colour naming scaled score	.890	**.004***
Colour-Word Interference, Word reading scaled score	.944	.097
Colour-Word Interference, Interference scaled score	.866	**.001***
Letter Fluency scaled score	.953	.173
Category Fluency scaled score	.966	.400
Trail Making Test A (s)	.772	**.000***
Trail Making Test B (s)	.579	**.000***
Judgement of Line Orientation (percentile)	.838	**.000***
Letter-number Sequencing scaled score	.976	.667
Block Design scaled score	.958	.236
Rey Complex Figure copy (ordinal score)	.672	**.000***
Rey Complex Figure immediate recall (T score)	.882	**.002***
Activities of Daily Living Scales		
Four-Item Instrumental Activities of Daily Living Scale	.970	.486
The Informant Questionnaire on Cognitive Decline in the Elderly	.915	**.015***
Alzheimer's dementia Activities of Daily Living International Scale	.951	.158

Table 9-9. Shapiro-Wilk *W* test scores for *SMCTests* measures for n=32 Dementia and Driving study

Test Measure	W Statistic	*p*-value (***p*<..05)**
Ballistic Movement Test		
Reaction time, right hand (ms)	.816	**.000***
Reaction time, left hand (ms)	.844	**.000***
Reaction time, grand mean (ms)	.885	**.003***
Movement time, right hand (ms)	.929	**.037***
Movement time, left hand (ms)	.819	**.000***
Movement time, grand mean (ms)	.920	**.021***
Total reaction and movement times, right hand (ms)	.905	**.008***
Total reaction and movement times, left hand (ms)	.916	**.016***
Total reaction and movement times, grand mean (ms)	.912	**.013***
Peak velocity, right hand (ms)	.933	**.048***
Peak velocity, left hand (ms)	.949	.133
Peak velocity, grand mean (ms)	.943	.091
Tracking Tests		
Sine tracking run 1 error (mm)	.914	**.014***
Sine tracking run 2 error (mm)	.731	**.000***
Random tracking run 1 error (mm)	.896	**.005***
Random tracking run 2 error (mm)	.846	**.000***
Arrows Perception Test		
Arrows Perception Test	.820	**.000***
Omission of arrows response	.436	**.000***
Divided Attention Test		
Tracking error (mm)	.965	.363
Number of arrows correct	.751	**.000***
Omission of arrows response	.436	**.000***
Complex Attention test		
Reaction time (ms)	.966	.403
Movement time (ms)	.968	.452
Total mean movement and reaction times (ms)	.982	.845
Reaction time standard deviation (ms)	.921	**.022***
Movement time standard deviation (ms)	.902	**.007***
Number of lapses errors	.578	**.000***
Number of invalid trials	.638	**.000***

Continued on following page

Continued from previous page		
Test Measure	W Statistic	*p*-value (***p*<..05*)*
Planning Test		
Lateral road position error (mm)	.755	**.000***
Duration of positional faults (s)	.740	**.000***
Distance travelled (m)	.817	**.000***
Intersection safety margin (mm)	.903	**.008***
Number of hazards hit	.851	**.000***
Number of crashes	.729	**.000***

11.20 Appendix T – Independent Measures Excluded from Offering to the n=32 Dementia and Driving Models with Reasons for Exclusion

Table 9-10. Measures that were excluded on pragmatic grounds and reasons for exclusion for the n=32 Dementia and Driving study

Measures not offered to BLR model	Reason for not offering
Handedness	Would not be used to predict in public health setting
Years of driving	Would not be used to predict in public health setting
Years of education	Would not be used to predict in public health setting
Forms of transport used regularly (e.g. car, passenger)	Descriptive only, would not be used to predict in public health setting
Wears glasses while driving	Descriptive only, would not be used to predict in public health setting
Self-rated quality of driving	Descriptive only, would not be used to predict in public health setting
Where the driver reports driving in the previous year (e.g. neighbouring and distant towns)	Descriptive only, would not be used to predict in public health setting
Geriatric Depression Scale	Effect in the opposite from expected direction (Fails had lower depression scores)
VOSP Silhouettes	Effect in the opposite from expected direction (Fails names more objects correctly)
Alzheimer's dementia Activities of Daily Living International Scale	Effect in the opposite from expected direction (Fails were rated by significant others as less impaired)
Ballistic Movement, Reaction time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Reaction time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Movement time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Total time, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Total time, left hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, right hand	Grand mean measure used instead of individual hand scores
Ballistic Movement, Peak velocity, left hand	Grand mean measure used instead of individual hand scores
Arrows Perception, number of arrows correct	Effect in the opposite from expected direction (Fails correctly reported the direction of the arrows more frequently)
Cancer	Effect in the opposite from expected direction (Fails less likely to report having cancer)
Higher blood pressure	Effect in the opposite from expected direction (Fails less likely to report having high blood pressure)
Depression	Effect in the opposite from expected direction (Fails less likely to report having depression)
Anxiety	Effect in the opposite from expected direction (Fails less likely to report having anxiety)
Falls	Effect in the opposite from expected direction (Fails less likely to report having falls)
Cataracts	Effect in the opposite from expected direction (Fails less likely to report having cataracts)