Using Serial Trichotomization with Neuropsychological Measures to Inform Clinical Decisions on Fitness to Drive Among Older Adults

Valerie Ranger

Thesis submitted to: School of Counselling, Psychotherapy and Spirituality
In partial fulfillment of the Masters in Psychotherapy
Supervisor: Stephanie Yamin, Ph.D., C. Psych.

School of Counselling, Psychotherapy & Spirituality
Saint Paul University

Abstract

For older adults, driving contributes to independence, autonomy, and wellbeing. For some older adults there comes a time when their ability to drive safely may be questioned due to illness or cognitive impairment. Decisions related to driving safety and when to cease driving are complex and often left to the clinical judgement of the primary care physician. There is an interest in developing a method that could help assist physicians in making that determination. To date, there is no well-defined cut-off point on neuropsychological tests that produces an acceptable level of sensitivity and specificity allowing for the determination of an individual's fitness to drive. Serial trichotomization involves classifying drivers as either pass, fail or indeterminate based on cut-points that lead to 100% sensitivity and specificity. The purpose of this study was to examine the serial trichotomization method using four common neuropsychological tests (i.e., 3MS, Trails A & B, clock drawing). Sensitivity and specificity for each test were established using a medical expert's clinical judgement. Charts of 105 patients at a tertiary memory disorders clinic were reviewed and data related to neuropsychological test scores and clinical judgement around fitness to drive were abstracted. After applying the serial trichotomization model, 38.1% of the sample were classified as unfit to drive, 36.1% were classified as indeterminate, and 25.8% were classified as fit to drive. This model could offer physicians a better preliminary method in assessing fitness to drive using common paper-pencil cognitive tests. With easy administration, a simple scoring system, and no equipment requirement, serial trichotomization allows primary care physicians to make specific and sensitive decisions around fitness to drive, in a timely fashion. This study adds to the growing body of literature supporting the use of serial trichotomization in making decisions about fitness to drive.

Acknowledgments

I would like to take the opportunity to thank a number of individuals who have helped me realize this thesis. Foremost, I would like to express my deepest appreciation to my thesis advisor Dr. Stephanie Yamin for her continuous support, knowledge, and encouragement that has greatly impacted my personal and professional growth. In addition, I would like to thank the rest of my thesis committee, Dr. Arne Stinchcombe and Dr. Cynthia Bilodeau for their insightful comments. I would like to thank Brenda Landry and Elizabeth Legace at the Memory Clinic for their kindness and generosity in assisting me throughout the recruitment process. I wish to recognize the support from Dr. Frank Knoefel and Dr. Andrew Frank for providing access to invaluable data. I would also like to thank my boyfriend Yvann for his continuous support, encouragement and editorial aid. Furthermore, I would also like to thank my family for their support, encouragement and understanding through this difficult and demanding process. Finally, I would like to thank Karen Molson for the review and editing of this thesis. Without the help of every one of you, this study could have not been possible and I am tremendously grateful.

Table of Contents

List of Tables

List of Figures

Using Serial Trichotomization with Neuropsychological Measures to Inform Clinical Decisions on Fitness to Drive Among Older Adults

The impact of the automobile on society

For most people in North America, the first step in going almost anywhere starts by reaching for your car keys. Car ownership has long had a profound impact on society and the lives of individuals, shaping the large-scale planning of urban infrastructure to the small-scale organization of one's work day. Owning a private vehicle provides flexibility and separates workplaces from homes, creating specific sectors. Even for those living in urban areas, owning a car has become essential to get to work, purchase food, attend appointments and take care of dependents. Most of these obligations need to be carried out every day in a short amount of time (Statistics Canada, 2007). Car dependency is especially high in rural areas where population density is lower; therefore, destinations are more dispersed and alternative transportation is limited or non-existent (Belton-Chevallier, Motte-Baumvol, Fol, & Jouffe, 2017). In 2016, it was estimated that four out of five Canadians used a private vehicle to get to work (Statistics Canada, 2011). Despite the growing accessibility of public transportation, the number of registered vehicles is still on the rise and has increased by 34.3 million from 2016 to 2017 (Statistics Canada, 2018). It appears that driving and having access to a private vehicle provide advantages that cannot be fulfilled by public transportation or other transportation options.

Smith, Shipley and Rose (1990) found that having access to a private vehicle is associated with a lower mortality rate after controlling for employment status and age. Car ownership is also associated with overall well-being and better health (MacIntyre, Hiscock, Kearns, & Ellaway, 2001). It is evident that the mere reality of owning a car does not have any impact on health, however, having better access to resources which the car provides can be a contributing factor to

better heath (Ellaway, MacIntyre, Hiscock, & Kearns, 2003). Researchers have also found that individuals who have access to a private vehicle gain psychosocial benefits. After controlling for age and socio-economic status, participants reported gaining protection, autonomy, prestige, and control by commuting with a private vehicle. Having the power and control over your environment has been shown to be related to well-being. Benefits appeared to be different according to gender, women gaining more protection and men gaining more autonomy (Ellaway et al., 2003). The association between car ownership and health can be better understood by the mediator relationship of psychosocial benefits, since it is well understood that psychosocial benefits contribute to better physical health outcomes and longevity (Brabdmaier, Ram, Wagner, & Gerstorf, 2017; Contrada et al., 2008; Lewebstein, 2002).

The Aging Population in Canada

The aging population is increasing in developed countries. For the first time in Canadian history, in 2016, adults over the age of 65 made up a bigger population segment than children under the age of 15 (Statistics Canada, 2016). There are a few reasons for this continuing trend. First, the generation of baby boomers born between 1946 and 1964 are getting older. Second, with the advancement of medicine, life expectancy is getting longer. Centenarians are the fastest growing segment of the population in Canada, increasing by 41.3% between 2011 and 2016. Lastly, lower fertility rates mean the population of children is decreasing. In 2016, approximately one in eight Canadians were aged 65 years and older. This segment of the population will continue to increase rapidly. It is estimated that in 2031, one in four will be over the age of 65.

It is essential to be aware of how aging trends affect specific geographic locations. Depending on the geographic location, the importance of particular needs will increase, notably the increased need for safe and accessible transportation (Statistics Canada, 2016). Concluding

from the 2006 census, the share of adults over the age of 65 living in predominantly rural areas was above the Canadian average. Older adults are attracted to rural landscapes in order to be able to retire and relax in spacious, quiet and natural environments (Statistics Canada, 2008). This attraction for rural living comes at a cost, with limited or inexistent public transportation options in rural areas, older adults become more car dependent and are forced to drive in order to fulfill their primary needs.

Aging and Driving

Since older adults constitute a growing segment of the population, there will be a corresponding increase of older drivers occupying the roads. Currently in Canada, three quarters of all older adults have an active driver's license and 200,000 are over the age of 85. In 2009, 85% of men aged 65 to 74 disclosed that driving their own vehicle remained their primary form of transportation. For Canadians over the age of 85, there is a big gap between male (67%) and female (26%) drivers. However, in the 45-65 age group, there are nearly as many women as men who drive and it is likely that these women will continue to drive into older age, therefore, the upcoming generation of drivers is expected to be even more numerous (Statistics Canada, 2012). According to the 2009 census, only 4% of adults over the age of 65 used public transportation; the proportion of individuals who used public transportation appeared to decline with age (Statistics Canada, 2012).

Statistics have shown that older drivers have the highest collision rate per kilometer after young male drivers (Statistics Canada, 2012). Statistics of the number of at-fault crashes by age group form a U-shaped curve with younger and older drivers having the highest rates of motor vehicle collisions. Young drivers have a high rate of collisions that slowly starts to decrease when they reach their mid-twenties, remains stable throughout their 30's to 50's, and begins to increase

again after the age of 65 (Ledger, Bennett, Chekaluk, & Batchelor, 2019). However, there are some differences between the crash patterns of younger and older drivers. Mayhew, Simpson, and Ferguson (2006), conducted a review on the type of collisions involving older adults in North America. They found that adults over the age of 65 are mostly involved in crashes during ideal driving conditions; clear days, straight roads, and dry pavement. Furthermore, they are more likely to be involved in multiple-vehicle, head-on, and angle crashes. Older adults are overinvolved in crashes resulting from changing lanes, merging, leaving a parked position, failure to yield the right of-way, and operating through intersections. A study conducted by Hauer (1988) found that 40% of traffic fatalities and 60% of injuries among older drivers occurred at intersections. Turning maneuvers, especially left turns at intersections appear to be problematic, as drivers need to understand the rules of right of-way and have a good perception of gap-acceptance between cars. Gap-acceptance refers to the available gap between one's own and an approaching vehicle. Intersections that have an extra layer of complexity increase the risk of collision for older adults. Intersections considered to be complex include those with no traffic lights, those that have more than two signals, and those with flashing lights (Mayhew et al., 2006). Younger drivers on the contrary, have an increased tendency to be involved in risk-taking collision including speeding, driving while under the influence of alcohol or drugs, and performing overtaking maneuvers. Young drivers are mostly involved in single-vehicle collisions during the night and in bad weather (Zhang, Fraser, Lindsay, Clarke, & Mao, 1998). Considering these crash patterns suggests that older drivers may have increased difficulties in perception, cognition, and inattention, since their collision rate is high even in ideal conditions.

Cognitive Abilities Involved in Driving

Driving is among one of the most challenging tasks performed daily. Multiple studies have found a relationship between cognitive functioning and driving abilities. As cognitive abilities decline, driving abilities have also been shown to decrease (Apolinario et al., 2009; Davis & Ohman, 2016; Fraade-Blanar et al., 2018; Hird, Vetivelu, Saposnik, & Schweizer, 2014; Ledger et al., 2019; Reger et al., 2004; Wagner, Muri, Nef, & Mosimann, 2011). Ledger and collegues (2019) have found that cognitive abilities could partially explain the similar crash rates among younger and older drivers. They found a relationship between cognitive functioning and driving performance for both age groups. In other words, lower cognitive abilities are associated with poor driving performance in both young and older drivers. Research has demonstrated that cognitive abilities decline with age, while some cognitive abilities are not fully developed when young adults start driving (e.g., prefrontal cortex and limbic system) (Sharma et al., 2013). Both groups that are statistically shown to be involved in the highest crash rates appear to have deficits in their levels of cognition. These findings highlight the implication of cognitive abilities in one's capacity to drive safely.

The research on specific cognitive abilities required to drive safely are determined largely by studying cognitively impaired individuals and noticing their precise driving difficulties. Researchers have found that certain cognitive domains are more closely related to driving performance; notably, mental status, executive functioning, visuospatial skills, attention, and memory (Anderson et al., 2012; Anstey, Wood, Lord, & Walker, 2005; Carr et al., 2011; Mathias & Lucas, 2009; Reger et al., 2004; Wagner et al., 2011). Mental status refers to the global and integral aspects of cognition. Executive functioning is a constellation of cognitive processes performed by the prefrontal regions of the brain and can be divided into three main categories:

working memory, inhibition, and set-shifting. Working memory enables individuals to manipulate and update information in the moment (Diamond, 2013). Inhibition refers to the ability to maintain attention of pertinent information while suppressing irrelevant stimuli. Set-shifting is associated with cognitive flexibility and refers to the capacity to rapidly change from one task to another while adapting to changing demands. Executive functioning is crucial to safe driving as drivers need to focus their attention on the road while keeping in mind their destination and route (i.e., working memory). Furthermore, they must continuously adjust their speed, follow traffic signs, be aware of other vehicles (i.e., set-shifting) and manage distractions from the scenery, and/or passengers (i.e., inhibition) (Walshe, McIntosh, Romer, & Winston, 2017).

Visuospatial skills include both the images we identify and our perception of the size, distance and location of our surroundings (Reger et al., 2004). This helps the driver position the vehicle accurately on the road, navigate turning maneuvers, change lanes and correctly park. Attention is the most extensively researched cognitive domain in relation to driving abilities and can be divided into three different types: selective, divided, and sustained (Reger et al., 2004). Selective attention is the capacity to focus on one thing at a time while divided attention is the capacity to focus on multiple things at once. Sustained attention is required to focus for long periods of time. Researchers have long debated what type of attention is best at predicting safe driving (Apolinario et al., 2009; Hird et al, 2014; Reger et al., 2004; Wagner et al., 2011). There are three categories of memory which are associated with safe driving abilities. Episodic memory refers to registering, acquiring and encoding information at a given time or place. For example, episodic memory is used by the driver to locate his keys. Semantic memory involves knowledge that an individual has accumulated throughout their lifetime. Semantic memory is critical in remembering the meaning of traffic signs and lights. Lastly, procedural memory is the ability to

remember learned skills that can be performed without devoting selective attention. Starting the engine and shifting are skills that require procedural memory and are important for safe driving (Wagner et al., 2011).

Cognition in Older Adults

Many studies in the field of cognition and aging have found that some cognitive abilities tend to decrease with age (Murman, 2015; Salthouse, 2009; Tuokko, Griffith, Simard, & Taler, 2016). However, it is important to distinguish the difference between normal age-related changes in cognition and changes that are associated with cognitive impairment due to medical conditions. Age alone does not negatively impact driving abilities, but rather medical conditions associated with aging do have an impact (Tuokko & Hunter, 2002). Evidently, there are serious consequences associated with allowing unfit drivers to drive, although, driving cessation has also been linked to multiple consequences: depression, reduced well-being (Aksan et al., 2015), nursing home placement, (Freeman, Gange, Muñoz, & West, 2006) and reduced community participation (Fairhall et al., 2014; Dahan-Oliel et al., 2010). Thus, among older drivers, it is important to find a balance between road safety and independence.

Normal cognitive decline is referred to as primary aging, meaning it is associated with an intrinsic process of biological aging despite good health and in the absence of medical conditions. Secondary aging suggests a deterioration that is age-related on a pathological level resulting from extrinsic factors like disease and medical conditions. They are classified as secondary aging effects since they are conditions that are more prevalent amongst the older population, although people can age without experiencing them. This is an important distinction to be made when examining questions related to aging and driving that are often overlooked (Tuokko & Hunter, 2002).

Cognition in primary aging. Fluid abilities and crystalized abilities are terms commonly used to describe cognitive changes related to aging and have different developmental trajectories. Researchers have found that fluid abilities tend to decrease with age while crystalized abilities increase with age. Fluid abilities are known to be responsible for processing and integrating new information, solving novel problems, and reflective manipulation. Crystalized abilities reflect the knowledge and skills acquired through education and life experiences. As individuals age, they have increased difficulty in learning new abilities and acquiring new knowledge (Murman, 2015). Furthermore, the size of the brain also decreases with age, a process referred to as atrophy (Harada, Love, & Triebel, 2013). For many years researchers assumed that the loss of gray matter volume was the result of neuronal loss; however with improved technology this has been shown to be untrue. Multiple studies have found that on average only 10% of neurons are lost when comparing older adults with young adults. To account for the change in gray matter volume, neurons do not die but rather change in structure (Pannese, 2011). The change in cognitive abilities for older adults is heterogeneous, meaning the rate of decline and the specific cognitive domains that are affected vary greatly from one individual to another (Murman, 2015). Due to those individual differences, the separation between primary aging and secondary aging can be difficult. The majority of healthy older adults do not pose an increased driving risk and are more likely to engage in self-regulatory driving behaviors such as limiting distances, remaining on well-known roads, and not driving at night or in bad weather (Harada et al., 2013). Normal cognitive changes associated with aging do not sufficiently account for higher collision rate in older drivers, especially for individuals below the age of 80 (Wagner et al., 2011; Evans, 2000). However, the onset of cognitive decline is subtle and oftentimes difficult to detect.

Cognition in secondary aging. Age-related disease can accelerate the normal aging process, thus, accelerating cognitive decline. Cognitive decline can be better understood when it is considered on a spectrum: normal aging, mild cognitive impairment, mild dementia, moderate dementia and severe dementia. Cognitive impairment can also be caused by several factors including extensive medication use, mental health problems, and a variety of medical conditions (Murman, 2015). The key differentiating factor between normal age-related changes in cognition and cognitive impairment is that in the latter case, the capacity to live independently is compromised and daily activities become difficult to accomplish alone. The number of Canadians living with cognitive impairment is currently estimated at 747,000 (Alzheimer Society of Canada, 2017). For the purpose of this review, we will focus on the most prominent conditions that are associated with cognitive impairment in older adults. According to the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5; American Psychiatric Association, 2013), cognitive impairments can be divided into two distinct categories: mild neurocognitive disorders and major neurocognitive disorders. Dementia is the leading cause of cognitive impairment in older adults and is considered to be a major neurocognitive disorder. However, the majority of patients first develop mild cognitive impairment, also known as a mild neurocognitive disorder.

Mild Cognitive Impairment (MCI). MCI is often regarded as an intermediate phase between normal expected changes in cognition due to aging and the development of dementia. The term was introduced by Reisberg and colleagues in the 1980's, however, the publication of international criteria for MCI only appeared in 1999 (Petersen et al., 2014). The difference between MCI and dementia can be better understood when considering the level of autonomy of a patient when examining their ability to perform activities of daily living (ADLs) and instrumental activities of daily living (IADLs). ADLs refer to the most basic self-care tasks that we learn as

young children which include walking, eating, dressing, bathing, and continence. IADLs are more complex tasks such as managing finances, driving, meal preparation, housecleaning, and communication. MCI is associated with subtle changes in IADLs while dementia leads to difficulty or inability in both IADLs and ADLs depending on the severity of progression (Gold, 2012). In a study conducted by Jutkowitz and colleagues (2015), IADLs changes were identified up to eight years before a diagnosis of dementia.

There are 3 distinct subtypes of MCI (Petersen et al., 2014). The first, amnestic MCI (aMCI) presents itself predominantly with memory impairment. The second, single-domain non-amnestic MCI, refers to the impairment of one of the following cognitive domains: language, visuospatial abilities, executive functions, or psychomotor abilities without experiencing decline in memory. Finally, multiple domain MCI includes decline in the aforementioned cognitive domains and can also include memory. Individuals with amnestic and multiple domain MCI are at an increased risk of developing Alzheimer's disease (AD). Although patients have a greater risk of developing dementia if they have a diagnosis of MCI, the progression is not certain. In a study conducted by Petersen and colleagues (2003), they found that the progression of MCI to dementia is approximately 12% per year and after 6 years, approximately 80% of MCI patients converted to dementia. Researchers have found that in some cases, MCI reverts to back to normal age-related cognition (Wada-Isoe al., 2012). The prevalence is variable from one study to another depending on the clinical criteria used in the diagnosis of MCI. MCI has not been associated with a specific brain lesion and many studies confirm a heterogeneous pathological profile between normal aging and dementia (Petersen et al., 2014). MCI is associated with various etiological subtypes (e.g., Alzheimer's, vascular, frontotemporal). A brief description of these subtypes will be elaborated upon subsequently.

Dementia. Dementia is an umbrella term for a large class of brain disorders characterized by the progression and irreversible deterioration of cognitive abilities. Dementia is best understood as a syndrome rather than a disease due to varied etiology. A syndrome refers to an agglomeration of symptoms that exist together to form a recognizable pattern (Dening & Sandilyan, 2015).

The ICD-10 Classification of Mental and Behavioral Disorders (World Health Organization, 1992) describes dementia according the following symptoms; decline in memory and thinking that interferes with everyday activities and care, impairment in registration, storage and retrieval of new or old information, reduction in flow of ideas, difficulty reasoning, and clear evidence of consciousness during early stages.

Dementia progression is frequently categorized in three main stages: mild, moderate and severe. Each phase of the syndrome is characterized by different psychological and behavioral symptoms that require different attention. In all forms of dementia, patients have difficulty or are unable to perform IADLs. Patients in the mild stages of dementia need very little or no assistance with ADLs and require very little supervision. However, in moderate phases, care is needed in most instances and patients typically struggle with ADLs. In this phase of the syndrome, patients start to have more pronounced behavioral changes including agitated behavior and exaggerated psychological effects; notably, hallucinations, delusions, and insomnia. In the last stage, patients are usually less disturbed and are incapable of accomplishing most ADLs and require constant care and attention. Furthermore, they are typically bedridden or wheelchair-bound and unable to move independently (Kua et al., 2014). The clinical dementia rating (CDR) scale is used to assess the severity of dementia patients. Through a semi-structured interview, patients are evaluated using a 5-point rating scale. Mild dementia is characterized by obtaining a score of 1, moderate dementia starting at 2 and severe dementia at 3. The scale is based on six domains of cognitive functioning

including memory, orientation, judgement and problem solving, community affairs, home and hobbies, and personal care (Morris, 1993). Dementia is a fatal disease, and death largely depends on the rate of progression, however, death usually occurs between three and 12 years after the diagnosis (Kua et al., 2014).

Age is the biggest risk factor in developing dementia but genetics and lifestyle during adulthood may modulate the development of the disease or syndrome (Sacuiu, 2016). Concurring from the Public Health Agency of Canada, the prevalence of dementia for adults between the age of 65 to 69 is 0.8%. The prevalence rises to 24.6% for adults over the age of 85 (Statistics Canada, 2017). Research has demonstrated that the prevalence of dementia is higher in women that it is in men. There are a number of possible factors that can contribute to this gender difference. One of the possible contributors is that women typically live longer than men, increasing their chances of developing dementia (Rocca, 2014). The incidence proportion is also rising. It is estimated that in 2038 there will be 257,811 new cases a year and 2.8% of the Canadian population will suffer from dementia (Alzheimer Society of Canada, 2010). Dementia has become a rising concern for Canadians and is now being referred to as an epidemic (Larson, Yaffe, & Langa 2013; Alzheimer Society of Canada, 2010; Sacuiu, 2016).

There are several different types of dementia, the most common being AD, vascular dementia (VaD), mixed dementia (MD), dementia with Lewy Bodies (DLB), and frontotemporal dementia (FTD). In addition, there are a number of medical conditions that can lead to the development of dementia, including Huntington's disease (HD), multiple sclerosis (MS), Parkinson's disease (PD), human immunodeficiency virus (HIV) (Dening & Sandilyan, 2015), traumatic brain injury (TBI) (Shively et al., 2012), Lyme disease (Kristoferitsch et al., 2018), and

drug related effects (e.g., chronic alcohol or psychotropic drugs) (Whishaw& Kolb, 2015). A brief description of the more prevalent etiological subtypes of dementia will be reviewed.

Alzheimer's disease (AD). AD is a progressive, irreversible, neurodegenerative disease of the brain impairing memory (Sacuis, 2016). It is the most common form of dementia and accounts for up to 75% of all cases (Qiu, Kivipelto, & Strauss 2009) and is the most common factor leading to cognitive impairment in older adults (Murman, 2015). AD was named after a German psychiatrist, Dr. Alois Alzheimer, who described the first case more than one hundred years ago. (Maurer, Volk, & Gerbaldo, 1997). AD is not part of the normal aging process as brain tissue degenerates causing decline in multiple cognitive domains including language, visuospatial and executive abilities. In the early stages, the most common clinical features are characterized by the loss of short-term memory and in later stages, orientation in place and time becomes impaired.

AD is characterized by two hallmark pathologies located in the brain causing cellular death. The first is referred to as *plaques* which are the deposits of beta amyloid protein that clump together in the brain. The plaques interrupt signals between neurons, ultimately causing cell death. The second is referred to as *neurofibrillary tangles* and are the collapses of a protein called Tau that ensures the transportation of nutrients and other substances to neurons. In a brain affected by AD, the protein twists and tangles, preventing nutrients to reach neurons and thus causing cell death. When there is a substantial amount of cell death in one region of the brain it causes atrophy, and in some cases patients can lose up to one-third brain volume (Whishaw& Kolb, 2015). Recently, there has been some debate on what causes these changes in the brains of patients afflicted by AD; however, no causal link has been found.

Vascular dementia (VaD). VaD is the second most common form of dementia after AD (Dening & Sandilyan, 2015; Kalaria, 2018; Korczyn, Vakhapova, & Grinberg, 2012; Murman,

2015; O'Brien & Thomas, 2015) and refers to a heterogeneous group of conditions resulting from different presentations of one or more cerebrovascular diseases causing dementia. Cerebrovascular disease occurs when there is reduced or blocked blood supply to the brain, depriving neurons of oxygen which inherently causes cellular death. Vascular dementia can be caused by different underlining diseases. The first, subcortical dementia is caused by small vessels in the brain that become twisted and stiff, reducing the blood flow to the brain (i.e small vessels disease). It is the most common condition that leads to VaD (Wardlaw, Smith, & Dichgans, 2013). Cerebral amyloid antipathy is similar to what happens in AD and can also be present in VaD. The accumulation of the amyloid protein causes plaque to build up in the blood vessels disturbing the normal blood supply to the brain. Other changes in the brain include infracts, or stroke and refer to the death of brain tissue when neurons get insufficient blood supply (Korczyn et al., 2012).

VaD's clinical symptoms are hard to decipher since vascular related brain lesions are heterogeneous, causing a variety of different cognitive impairments depending on the affected brain regions (Korczn et al., 2012). The most common expressions of VaD include decline in executive functioning, impaired attention, changes in the ability to make decisions, plan, or organize, and difficulties with movement such as walking and balance (Nordlund et al., 2007). Symptoms tend to have a slow onset and progress in stages, meaning that abilities deteriorate, stabilize, then deteriorate further. Cognitive impairment may rapidly increase after a stroke.

Mixed dementia (MD). MD includes clinical characteristics and brain changes of both AD and VaD. Despite them being recognized as distinctive types of dementia, some patients experience degenerative effects of AD and vascular ischemia (i.e., reduced blood flow to the brain). It was first described in 1962 by Delay and colleagues and clear diagnostic criteria are still being debated. It is difficult to estimate the prevalence of MD as most cases are only apparent

during post-mortem examination and are likely underestimated. The diagnostic criteria for AD and VaD often overlap and recent studies suggest that pure AD and VaD are unlikely (Zeky, Duyckaerts, & Hauw, 2007).

Dementia with Lewy Bodies (DLB). DLB is the third most common type of dementia (Sonnen et al., 2007). It is a neurodegenerative disease that occurs due to the accumulation of alpha-synuclien protein referred to as Lewy bodies and named after Fritz Lewy in 1912. Lewy bodies are found in the brainstem nuclei, limbic system, parahippocampal cortices, amygdala, and cortex. The deposits of this protein can also be found in Parkinson's disease (PD), and to a lesser extent in AD. Parkinson disease dementia (PDD) is also characterized by diffuse Lewy bodies making the distinction difficult (Mayo & Bordelon, 2014). Furthermore, DLB can also have similar brain deterioration found in AD, notably, plaques and neurofibrillary tangles. The cause of Lewy bodies in DLB remains unclear.

DLB core clinical features are characterized by fluctuation in cognition, attention and arousal. Visual hallucinations are a distinguishing clinical feature of DLB and occur in 80% of patients. Furthermore, DLB is commonly accompanied by Parkinsonism that includes bradykinesia (i.e., slowness of movement), rest tremor and or rigidity. Finally, DLB patients can experience REM sleep behavior disorders which may be present before cognitive decline (McKeith et al., 2017). DLB is more commonly diagnosed in men and progression of DLB is usually rapid. Confirmation of diagnosis can only be confirmed at the time of an autopsy (Mayo & Bordelon, 2014).

Frontotemporal dementia (FTD). FTD is an umbrella term that includes a group of rare neurodegenerative diseases affecting the frontal and temporal cortices. The first description was made by Alois Pick in 1892 and in 1911 Alois Alzheimer made the parallel between the

characteristics associated with Pick bodies and named it Pick's disease. Pick's disease is used as a synonym for FTD and refers to Pick bodies found in neurons making them increase in size. The main features of FTD are progressive degeneration in behaviour, executive function, or language (Bang, Spina, & Miller, 2015).

FTD is classified into three sub-groups: behavioural-variant frontotemporal dementia (bvFTD), non-fluent variant primary progressive aphasia (nfvPPA), and semantic-variant primary progressive aphasia (svPPA). The behavioral variant is responsible for approximately half the cases of FTD, is characterized by progressive declines in social functioning, and includes personality changes. The two remaining subtypes are associated with aphasia, an impairment of language. NfvPPA is associated with motor-speech difficulties and agrammatism, (i.e., the inability to communicate in full sentences using functional words), while svPPA's clinical features includes the loss of semantic knowledge referring to the meaning of words (Bott, Radke, Stephens, Kramer, 2014). FTD is the leading type of dementia for patients bellow the age of 65 (i.e., early-onset dementia). Diagnosis of FTD is estimated to be under-reported as behavioural changes are closely related with behavioural changes that occur in psychiatric disorders (Olney, Spina, Miller, 2017).

Cognitive Impairment and Driving

In 2009, 28% of Canadians diagnosed with a form of dementia had active drivers' licenses and 7,000 of these licences were held by patients with advanced stage dementia (Statistics Canada, 2012). As mentioned previously, MCI and dementia patients can experience a wide range of cognitive impairments important in one's ability to safely operate a vehicle. Patients who are diagnosed with dementia are between two and five times more at risk of being involved in a collision (Davis et al., 2018), yet a diagnosis of dementia does not automatically trigger the

cessation of driving for a patient. As Canada moves towards a concept of person-centered care, where individuals are evaluated based on their own abilities and limitations, driving with cognitive impairment becomes a complex issue to assess (Apolinario et al., 2009).

Recently, research on cognitive impairment and driving have recognized the importance of acknowledging the specific characteristics of different types of this diagnosis in order to get a better understanding of the risk factors associated with safe driving (Wagner et al., 2011). Many studies have found that dementia severity is the main factor in driving cessation. Increased ratings on the Clinical Dementia Rating (CDR) scale is associated with a higher risk of collision and impaired driving (Wagner et al., 2011). Moderate to severe dementia cases are often a clear determination that the patient is unsafe to drive; however, making recommendations on milder cases remains challenging. Research examining the implications of MCI and driving is still in its infancy. Studies have found that some individuals with MCI have minor differences in their left-hand turns, and lane control abilities. It was noted that in their overall rating, when compared to control groups of equal age, differences were not sufficient to account for a significant effect (Frittelli et al., 2009; Wadley et al., 2009).

Multiple researchers have found that some patients with dementia have the necessary abilities to drive safely for a certain period of time if they undergo testing regularly (Brown et al., 2005; Duchek et al., 2003; Fox, Bowden, Bashford, & Smith, 1997; Ott et al., 2008). Herrmann and colleagues (2006) conducted a study in collaboration with the Canadian Outcomes Study in Dementia (COSID) and found that 51.5% of mildly demented patients were still able to drive safely after 2 years. Similar studies have found comparable results (Adler & Kuskowski, 2003; Foley, Kamal, & Masaki, 2000).

Chee and colleagues (2017) conducted an international meta-analysis on the risk of motor vehicle collisions in relation to driving impairment among dementia patients. They highlighted that the majority of studies found decreased performance in at least one measure of driving behavior when compared to healthy control groups. Landmark/sign identification errors (i.e., verbally identifying road signs) and number of lost trips (i.e., unable to memorize a 4-turn route) were most commonly found in drivers diagnosed with dementia. Furthermore, they found that drivers diagnosed with dementia were at an increased risk of failing an on-road driving assessment. At the same time, two studies in the meta-analysis found that patients with mild dementia achieved passing or marginal scores on an on-road driving assessment. The researchers have pointed out that isolating true cases of mild dementia can be problematic due to its similarity to MCI. Consensus has been reached that identifying driving concerns should be raised in the earlier stages of the disease when patients are diagnosed with MCI or mild dementia as they may start to be less aware of their limitations, may be involved in more risk-taking behaviour and less likely to engage in self-regulatory driving behaviours. A limited number of studies have been carried out to evaluate the driving performance of patients with moderate dementia due to the fact that very few patients at this stage are able to operate a vehicle. The limited amount of data available implies that drivers with moderate dementia fail the most basic on-road driving assessments (Berndt, Clark, & May, 2008; Berndt, May, & Darzins, 2015).

Cognitive decline and progression differ greatly in each patient. Some patients' driving abilities will be affected more rapidly than others. MCI and mild dementia are critical stages of the disease where the individual's driving abilities should be evaluated closely as some are fit to drive while others are not. However, in the life of every dementia patient there will come a time when they will no longer be able to drive. This represents a significant loss in their lives, more

than a loss of independence and a perception of being a burden to their family, but it also has other negative effects on their self-esteem and identity including depression and isolation (Berndt et al., 2015). Thus, the assessments surrounding older adults' fitness to drive must be made with accuracy and precision.

Responsibility of Assessing Fitness to Drive

Currently, there are various regulations across provinces in Canada on licence-renewal procedures for older adults because responsibility is enforced under the provincial or territorial authorities. A scan of provincial policies indicates that neither Saskatchewan, Manitoba, New Brunswick, Nova Scotia, Prince Edward Island, Yukon, Northwest Territories, nor Nunavut have any license renewal requirement based on age. British Columbia, Alberta, Quebec, and Newfoundland and Labrador require older adults to provide a medical report from their physician in order to have their licence renewed. Ontario has a licence renewal program that consists of a vision test, a multiple-choice driving aptitude test, a group education session, a driver record review, and two brief cognitive screening tests (i.e. Clock Drawing test and Letter Cancellation Test). In some cases, they may be required to take an on-road test (National Association of Federal Retirees; Ontario Ministry of transportation, 2017). Moreover, most jurisdictions impose on physicians, optometrists and registered psychologists a duty to report patients whom they deem unfit to drive (Tuokko & Hunter, 2002). The decision related to driving safety and when to cease driving are complex and often left to the clinical judgement of the primary care physician. *Determining Medical Fitness to Drive: A Guide for Physicians* (2000) is a book published by The Canadian Medical Association to help inform physicians on conditions that can affect driving. Unfortunately, this guide book does not outline a specific protocol to help identify areas that require assessment. According to a study done by Wilson and Kirby (2008) there is inconsistency

in the content and procedures of assessments done by physicians, especially when assessing cognitive functions. Only 59% of physicians addressed the subject of driving, with their patients diagnosed with dementia (Davis et al., 2016). Many physicians (45%) are not confident in this subject area and do not consider themselves qualified to determine whether a patient is unfit to drive. About 85% of general practicioners expressed that they could benefit from further education in assessing driving fitness. Although most physicians agree that they should hold legal responsibility in reporting unsafe drivers (Jang et al., 2007), there is still an absence of rigorous guidelines to support clinical judgement (Dickerson et al., 2017; Herrmann, 2006).

Off-Road Assessment of Driving Fitness

There is recent interest in developing an off-road test to help physicians better determine fitness to drive amongst older adults. Furthermore, physicians are imposed a duty to report patients who are unfit to drive at any time without much guidance. An off-road driving test could optimize efficiency, provide a safer way to assess fitness to drive (Gibbons et al., 2017) and facilitate routine assessments (Unsworth et al., 2018). Neuropsychological tests have shown potential in being a great choice for an off-road test for primary care physicians as they typically require limited resources, equipment, and simple and short administration. Over the years, there have been many attempts to develop a valid off-road test utilizing neuropsychological tests without much success (Dickerson et al., 2017; Kay et al., 2011; Reger et al, 2004).

Neuropsychological tests to assess fitness to drive. Neuropsychological tests are tests that are designed to measure specific cognitive domains that are known to be linked with a particular brain pathway or structure. These tests use a bottom-up approach to assess the brain's function to gain insight on what specific cognitive domains are impacted in the patient. The profile of each patient who experiences cognitive impairment is heterogeneous despite sharing similar

etiological subtypes. This offers physicians an accurate depiction of specific cognitive domains that may be impacted in each individual patient (Burns et al., 2018). Driving is among one of the most challenging IADLs performed daily and requires the integration of multiple cognitive domains; notably, mental status, executive functioning, visuospatial skills, attention, and memory (Anderson et al., 2012; Anstey, Wood, Lord, & Walker, 2005; Carr et al., 2011; Mathias & Lucas, 2009; Reger et al., 2004; Wagner et al., 2011). Many researchers have reported the predictive value of neuropsychological tests to evaluate the driving performance of older adults who experience cognitive impairment (Anderson et al., 2014; Anstey et al., 2005; Apolinario et al., 2009; Hird et al., 2014; Mathias et al., 2009; Ledger et al., 2019).

Reger and colleagues (2004) conducted a meta-analysis of 27 studies to examine the relationship between neuropsychological tests and driving ability in dementia patients. The studies used on-road driving assessments and simulator tests to find an association between neuropsychological tests and fitness to drive. The findings revealed that both the on-road driving assessment and simulator test correlated with general mental status and attention, while the simulator test also independently correlated with visuospatial skills, executive function and memory. Multiple neuropsychological tests were used to evaluate each cognitive domain. The Mini-Mental State Examination (MMSE) was used to assess general mental status while other studies have also used the Montreal Cognitive Assessment (MoCA) (Burns et al., 2018; Esser et al., 2016; Gibbons et al., 2017). Attention was evaluated using a number of different tests including Part A of the Trail Making Test, the Digit Span Test, the Useful Field of Vision Test, the Letter Cancellation Test, and the Visual Tracking Test. Furthermore, the Benton Copy Test, the Clock Drawing Test, the Figure-Ground Test, the Hooper Visual Organization Test, the Mattis Construction Test, and the Visuospatial Task of the Stanford-Binet Intelligence Test were used to

evaluate drivers' visuospatial skills. Memory was evaluated using the following neuropsychological tests: the Logical Memory Test, the Benton Visual Retention Test, the Visual Reproduction Test, and the Facial Recognition Test. Finally, Part B of the Trail Making Test, the Word Fluency test, the Stroop Color-Word Test, the Picture Arrangement Test, and the Category Fluency Test were used to assess drivers' executive functions. The study highlighted the importance of neuropsychological tests in predicting driving abilities in older adults with cognitive impairments. However, they stressed the importance of further research in establishing appropriate cut-off points on neuropsychological tests to indicate at what level a patient is unfit to drive.

Despite the fact that researchers have found a significant association between neuropsychological tests and on-road driving performance in older adults with cognitive impairments, these tools can only be used to trigger the need for further evaluation. This is because reliable and valid cut-off points have not yet been established (Dickerson et al., 2017; Kay et al., 2011; Reger et al, 2004). Cut-off points are used to categorize drivers as either fit or unfit based on a specific score they obtain on the neuropsychological tests. Misclassification errors could have a substantial impact on a person's life, therefore, off-road tests need to be rigorously examined for their validity.

Reference standard used to determine validity. If off-road driving tests are to be used to assess fitness to drive, they must be valid, thus truly measuring a person's ability to operate an automobile. In most cases, off-road tests are validated against a reference standard that can include on-road driving tests, crash/violation statistics, and simulators. Every reference standard has its own advantages and disadvantages that can impact the quality of the evidence generated for predicting driving capacity (Kay et al., 2012).

On-road driving test. The gold standard to assess fitness to drive remains an on-road driving test as it demonstrates high face validity (Gibbons et al., 2017; Hoggarth, Innes, Dalrymple-Alford, Severinsen, & Jones., 2010; Kay et al., 2012; Shechtman, Awadzi, Classen, Lanford, & Joo, 2010). However, there is no standardized or widely used on-road driving test that is applied within the provinces, or across Canada (Shechtman et al., 2010). Standardization is very difficult to achieve using an on-road test since driving routes and manoeuvres are dependent on the area in which the test is conducted (Kay et al., 2012; Schechtman et al., 2010). Older adults may not perform optimally due to anxiety associated with performing the on-road driving test. Furthermore, on-road driving tests lack efficiency from a time perspective (Kay et al., 2012). Older adults are put on a waiting list and are usually required to wait many weeks or months before being evaluated. This method of evaluation is costly and dangerous as some older adults have impaired abilities that can endanger themselves, the assessors, and other drivers on the road.

Crash/ violation statistics. Crash/violation statistics can be collected by official records or by self-reporting. These statistics represent potentially unsafe driving occurrences and may identify unfit drivers (Kay et al. 2012). Statistics collected by official records demonstrate objectivity, however they are not the most reliable measure since not all unfit drivers will be involved in a collision. Older adults may have close encounters which are not being accounted for using this reference standard. Furthermore, not all incidents are officially recorded (Lew, Poole, Lee, Jaffe, Huang, & Brodd., 2005) and self-reporting also lacks reliability as it is subject to the social desirability bias.

Simulators. Driving simulators have the advantage of standard driving routes and manoeuvres performed in a laboratory setting without endangering other citizens (Kay et al., 2012) but there is a lack of consistency on whether simulator-based assessments correspond with on-road

driving abilities. Other studies have found that reaction time on driving simulators does not adequately represent on-road skills (Lew et al., 2005). Moreover, some older adults may experience simulator sickness. According to a study conducted in various institutions by Tick and Caird (2011), utilizing a simulator to assess older adults' fitness to drive led, on average, to a 40% dropout rate due to simulator sickness. Symptoms related to simulator sickness include general discomfort, fatigue, eye-strain, blurred vision, headache, difficulty focusing, dizziness, vertigo, burping, sweating, nausea, and increase salivation. Typically, participants who drop out experience stronger symptoms (Matas, Nettelbeck, & Burns, 2015).

Experts' clinical decision. Clinical judgment has not yet been used as a reference standard to validate off-road tests as the majority of physicians are not confident in their ability to assess fitness to drive (Jang et al., 2007; Marshall, Demmings, Woolnough, Salim, & Man-Son-Hing, 2012). However, some physicians, neuropsychologists and occupational therapists have developed expertise in the area of fitness to drive in older age. Their expertise is often drawn from numerous years dedicated to driving research, practice, and consultation on issues related to driving fitness and aging. These professionals have gained years of knowledge on the topic and are equipped to addresses the complex and ambiguous questions related to fitness to drive. The National Older Driver Research and Training Center recognized the importance of experts' opinions regarding assessing fitness to drive and held the International Older Driver Consensus Conference where 63 international experts met to discuss criteria for determining whether an individual is fit or unfit to drive and where more work is needed to better assess fitness to drive (Stephens et al., 2005). Experts' predictions about a demented patient's ability to drive safely has been reported to demonstrate validity. A study conducted by Brown and colleagues (2005) compared the predictions of driving performance by dementia patients, family members, and an experienced

neurologist and dementia specialist for 50 participants with AD using a standardized road test as the criterion. The expert's prediction of fitness to drive was significantly predictive of on-road driving performance with a 90.0 sensitivity and 60.7 specificity. Fox and colleagues (1997) found similar results in that a neuropsychologist's predictions on AD patient driving abilities were found to be significantly associated with on-road driving performance. However, similarly to the on-road driving test, clinical experts do not have a standardized or widely used protocol or tests to assess fitness to drive, though a high level of control can be obtained when the same health professional or inter-judge agreement is used to asses all participants. Experts' clinical judgments require far fewer resources than the aforementioned reference standard and provides a time efficient way to evaluate fitness to drive amongst older adults.

Performance of a classification test. Correlation or regression models are often used to demonstrate the significant association between driving performance (e.g. on-road driving tests, crash/violation statistics, simulators, experts' clinical judgements) and off-road test results. Researchers have argued that a statistically significant association between an off-road test and its reference standard are preliminary steps, although not sufficient. Basic screening test properties must be reported to truly understand their validity (i.e., the extent that the test measures what it is designed to measure, in this case fitness to drive).

There are 4 basic properties of a classification test; (1) sensitivity, (2) specificity, (3) positive predictive value (PPV), and (4) negative predictive values (PPN) (Weaver, Walter, & Bédard, 2013). These properties will be explained in more detail in the following paragraph.

Fitness to drive is a binary decision; an individual is either fit or unfit to drive. However most cognitive tests are not binary, but rather, they are continuous and are dependent on a score. For example, the Clock Drawing scores typically ranges from 0-7. Thus, a cut-off point needs to

be established to categorize individuals who scored above that point as fit and categorize individuals below that point as unfit. Performance of a classification test arises from the placement of the cut-off point. (1) Sensitivity and (2) specificity are used to measure if the cut-off point has successfully placed individuals in the appropriate category. (1) Sensitivity identifies true positives: individuals who are unfit to drive and are correctly identified as such by the test. (2) Specificity identifies true negatives, who are fit drivers who are correctly identified as such by the test. The difficulties in preventing false positives are associated with low specificity; thus there is an increased risk that safe drivers will be labeled as unfit by the test. Difficulties in preventing false negatives are associated with low sensitivity that may inherently result in an increase number of misclassification errors that allow unfit drivers to continue to operate a vehicle. See Figure 1 for a visual representation. (3) Positive and (4) negative predictive values are used to determine the probability that the established cut-off points will categorize drivers correctly. Positive predictive values determine how likely it is that an individual who is unfit to drive — as determined by an off-road test — is truly unfit. On the contrary, negative predictive values refers to the probability that individuals who are determined fit to drive by the off-road test are truly fit to drive (Weaver et al., 2013).

In a perfect classification test, sensitivity and specificity would be equal to 100 percent respectively and positive and negative predictive values would also be equal to 100 percent. Given the importance of correctly identifying whether a person is fit to drive, it is imperative that an off-road driving test be highly sensitive and highly specific in order to avoid misclassification errors that can have serious impacts.

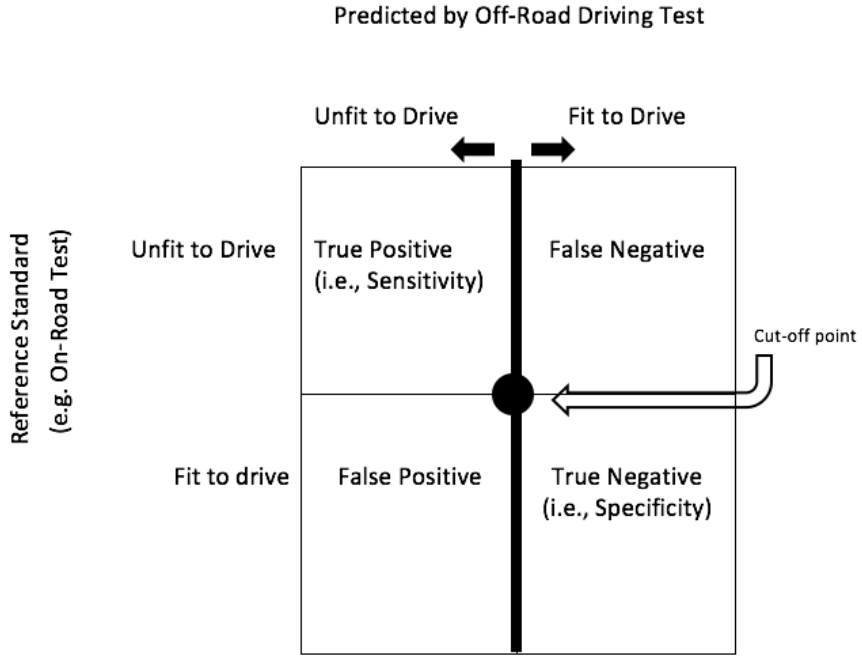


Figure 1.
Standard layout of 2X2 table modified to include basic properties of a classification test used to assess fitness to drive.

Binary classification systems. A panel of experts reached consensus on the following statement: "A decision about continued, restricted, or cessation of driving should never be made on the results of one tool in isolation, as there is not enough evidence provided by any tool to make a decision" (Bédard & Dickerson, 2014, p.128). Several studies have drawn similar conclusions (Carr, Barco, Wallendorf, Snellgrove, & Ott, 2011; Dickerson et al., 2017; Gibbons et al., 2017). There are three main reasons why one single test or assessment tool is not adequate to evaluate fitness to drive. First, driving involves a combination of multiple skills (e.g., attention, visual scanning, reaction time, etc.), a single test or screening tool may not evaluate all abilities required for driving. Second, all tests and screening tools have their inherent limitations. Using only one test or screening tool increases the risk that that these limitations are exacerbated during the assessment. Lastly, tools and standardized tests can only be used in conditions under which they

have been previously tested (i.e., a test that has yielded results for participants who have traumatic brain injury cannot be generalised for participants who have had a stroke) (Bédard & Dickerson, 2014).

Multiple tests. Researchers have then tried to combine multiple off-road tests together to increase precision. Aslaksen, Ørbo, Elvestad, Schäfer, & Anke (2013) determined that prediction of driving ability after traumatic brain injury and stroke can be best assessed using the California Computerized Assessment Package (CalCap), the Trail Making Test Part A (TMT A), and the Grooved Pegboard test. The CalCap measured visuomotor reaction time with 77% sensitivity and 23% specificity. Cognitive/psychomotor speed was measured using the TMT A with 85% sensitivity and 28% specificity. Finally, the Grooved Pegboard test demonstrated 82% sensitivity and 29% specificity while measuring visuomotor/fine motor speed.

Similarly, Carr and colleagues (2011) determined that three tests (the Snellgrove Maze Task, the Eight-Item Informant Interview to Differentiate Aging and Dementia, and the Clock Drawing Task) were the best at predicting road test performance in drivers with dementia, although, these tests only accounted for 24% specificity and 98% sensitivity. Wood, Anstey, Kerr, Lacherez, and Lord (2008) used a battery of tests that assessed visual, cognitive, and sensorimotor domains for predicting fitness to drive in older adults. The final model which included color choice reaction time, number of kilometers driven, postural sway, and motion sensitivity achieved 91% sensitivity and 70% specificity. According to a critical review of off-road tests predicting on-road performance done by Kay, Bundy, Clemson, Cheal, and Glendenning (2009) the Computerised Sensorimotor and Cognitive Test yielded the best combination of sensitivity and specificity with both measures exceeding 90% (Innes et al., 2005). However, when the cut-off points were applied to an independent sample, sensitivity and specificity dropped remarkably.

Off-road driving tests using a binary classification system to assess fitness to drive have demonstrated to be poor measures of performance (i.e., sensitivity and specificity). The aforementioned systems lack the capabilities of preventing false positives and false negatives where safe drivers are labelled as unfit (i.e., unfair to the driver) or unsafe drivers are labelled as fit (i.e., safety risk). Multiple researchers have expressed their difficulties in establishing appropriate cut-off points, from which stems sensitivity and specificity (Carr et al., 2011; Gibbons et al., 2017; Kay et al., 2012). As shown in the previous binary classification systems, it is very difficult to maintain a high level of specificity while maintaining a high level of sensitivity. There is generally a trade-off between both statistical measures; as specificity goes up, sensitivity goes down and vice-versa (Kay et al., 2012). There is no universal rule on an acceptable percentage of specificity and sensitivity. Acceptance depends on the seriousness of making the wrong decision. In the case of off-road driving tests, higher sensitivity is prioritized because incorrectly labeling a safe driver as an unfit driver seems a better outcome than labeling an unsafe driver as fit.

Trichotomization. Trichotomization is the act of classifying something or someone into three separately defined categories (Webster, 1963). Recognizing the difficulty of finding appropriate cut-off points to establish a highly sensitive and a highly specific test to assess fitness to drive, some researchers have trichotomized drivers into three distinct categories (e.g., unsafe, safe, and further testing). As mentioned previously, there is a trade-off between sensitivity and specificity as one cut-off point must try and optimize both of these performance measures. Trichotomization enables the establishment of two separate cut-off points. The upper cut-off point maintains a high level of specificity while the lower cut-off point ensures a high level of sensitivity. This creates a gap between the upper cut-off point and the lower cut-off point where participants are nether categorized as fit or unfit to drive therefore needing to undergo further testing.

Kay, Bundy, & Clemson. (2009), trichotomized drivers into unsafe, safe, and further testing categories. The purpose of their study was to examine the predictive validity of DriveSafe and DriveAware Tests (i.e., test measuring drivers' awareness of the driving environment and of driving abilities) when compared to the performance of an on-road driving test by using an optimal lower cut-off point to assess unfit drivers and an optimal upper cut-off point to determine fit drivers. Fifty percent of participants whose test scores were found between the upper cut-off point and the lower cut-off point required an on-road driving test and were categorized in the further testing group. The test achieved 96% to 97% specificity and 93% to 95% sensitivity. The high specificity and sensitivity showed that the use of a trichotomization classification system over a binary classification system could be a possible solution in improving the validity of off-road driving tests.

Dobbs and Schopflocher (2010) created a tool, Screen for the Identification of Cognitively Impaired Medically At-Risk Drivers (SIMARD), and used trichotomization in order to categorize individuals in three distinct categories: safe, unsafe and indeterminate. The SIMARD is a test that includes a series of tasks, involving word list, number transcoding, semantic world fluency task, digit span backward, and delayed recall. The test yields 86% sensitivity and 84% specificity. Forty-nine percent of participants were categorised as indeterminate. Wernham and colleagues (2014), conducted a study to compare the clinical decision made by a geriatrician regarding fitness to drive with the SIMARD and the suggested cut-off point. The study found no association regarding the decision to assess fitness to drive between the SIMARD and the clinical decision with patients diagnosed with mild dementia and mild cognitive impairment. Other studies have also debated the validity of SIMARD (Bédard et al., 2013). As stated previously, the use of one

single test is not sufficient in predicting fitness to drive as it results in a large grey area (i.e., participants require further testing) (Bédard & Dickerson, 2014).

Serial trichotomization. Similarly, Gibbons and colleagues (2017) have used serial trichotomization to assess fitness to drive, with more precision when using five common neuropsychological tests: the Trail Making Test Part A and Part B, the Montreal Cognitive Assessment, the Motor-Free Visual Perception Test (Third Edition) and the Clock Drawing Test. The main purpose of their study was to establish upper and lower cut-off points for each for the five cognitive tests reflecting 100% sensitivity and 100% specificity using results of an on-road driving test as a reference standard. Upper cut-off points were used to determine fit drivers and lower cut-off points were used to evaluate unfit drivers. These cut-off points were established using the coordinate points of the receiver operating characteristic (ROC) curve performed for each neuropsychological test. Once cut-off points were established for all five cognitive tests, participants were rated based on the first neuropsychological test, as pass, fail or indeterminate. Drivers rated as indeterminate were funnelled through to the next test. This rating continued until the indeterminate individuals were funnelled through all five tests. At the end of this process only 21.7% of all drivers in the trichotomization classification system remained indeterminate and needed further testing. This approach can be done with routine tests and demonstrates having great validity. However, serial trichotomization to assess fitness to drive cannot yet be seen as definitive, as it has only ever been tested once with a relatively small sample size (n=83). Furthermore, cut-off points were established using a sample of participants with a heterogeneous clinical profile. Some participants did not have any obvious clinical profile of cognitive impairment (i.e., physical injury, chronic pain, spinal injury, or amputation) while others had a more obvious clinical profile of cognitive impairment (i.e., cerebrovascular accident, traumatic brain injury, or cognitive

deficits). Moreover, an on-road driving test was used as the reference standard and could have potentially impacted the quality of the evidence generated for predicting driving capacity. Despite the fact that the on-road driving test is considered to be the gold-standard to validate the off-road driving test, they are prone to lack of standardization and the inability to control certain variables (e.g., traffic flow, environmental conditions, and other drivers) (Kay et al., 2012; Reger et al., 2004). Thus, a different clinical setting and different reference standard could possibly yield different cut-off points.

The aim of this study is to replicate the serial trichotomization method of Gibbons and colleagues (2017) in a different clinical setting using experts' clinical judgement as a reference standard indicative of driving fitness. Older adults with cognitive impairments will represent the sample of participants. It is expected that this study will draw similar cut-off points to the findings of Gibbons and colleagues (2017). Furthermore, with a bigger sample size, this study could reevaluate the proportion of older adults who could be categorized as either fit or unfit to drive without requiring further examination. This study would add to the growing body of literature supporting the use of serial trichotomization to make decisions around fitness to drive and to help create valid and reliable cut-off points on neuropsychological tests used for routine driving assessments.

Method

Research Design

This study used a retrospective chart review to analyze files of patients who were seen at the tertiary memory disorders clinic at Elisabeth Bruyere Hospital in Ottawa, Ontario. Ethics approval was obtained for this study from the research ethics boards (REB) at Saint-Paul University and Bruyere Research Institute.

Participants

Through the collaboration of two expert physicians, test scores from 109 patients (n=109) seen at the Bruyere Memory program at the Elisabeth Bruyere Hospital were abstracted. Four participants were excluded due to recurrent seizures and significant apraxia. According to the Canadian Council for Motor Transport Administrators (CCMTA) Medical Standards for Drivers (2017), individuals are prone to license suspension if seizures are unpredictable and unstable. Secondly, significant apraxia can disable participants' fine motor skills (e.g., writing) and can prevent participants from completing certain neuropsychological tests (e.g., Clock Drawing Test). Based on these premises, the four aforementioned participants were excluded from the study. The final sample size was comprised of 105 participants (n=105). The sample size was justified on the grounds that the previous study using the serial trichotomiztion method (Gibbons, 2017) had reported a smaller sample size (n=83). One of the aims of this study was to replicate the serial trichotomization method using a bigger sample size.

A convenience sample of medical charts documenting patients who had experienced cognitive impairment were selected for review. Only medical charts indicating patients' acceptance to participate in research (i.e., signed agreement found in the chart) were considered. Only medical charts who met inclusion and exclusion criteria were included in the analysis (refer to table 1). All participants were above the age of 50 and they completed cognitive tests in English. The mean age was 70.7 (SD= 9.69) years of age with a range of 53-90. The sample was comprised of 42.9% females and 57.1% males.

Table 1.
Inclusion and Exclusion Criteria used for Recruitment.

Inclusion	Exclusion
• Patients above the age of 50.	• Patients that have a mood disorder or significant anxiety disorder diagnosis.
• Patients who have experienced cognitive impairment.	• Patients who have an active substance abuse/misuse diagnosis.
• Patients who had an active driver's license at the time of their first consultation at the Bruyere Memory program.	• Patients who have a physical impairment which may impact their fitness to drive.
• Patients of the two chosen expert physicians (for control purposes as a determination of fitness to drive must be documented).	• Patients that have completed cognitive tests in languages other than English.
• Patients who were seen at the Bruyere Memory program at least 6 months from the data collection point (to allow for processing of the file and determination of fitness to drive by the physician).	• Patients who have not completed all of the cognitive tests (3MS, Trails A &B, and clock drawing). • Patients who completed abbreviated versions of the tests were excluded

Data Sources and Neuropsychological Measures

Demographic data. Demographic information was abstracted from charts when they met inclusion and exclusion criteria. Data were gathered concerning age, sex and medical issues that precipitated referral.

Trail Making Test. The trail making test is comprised of two parts; Part A and Part B. For the purpose of this study, both parts were used. The tasks require participants to connect a series

of 25 circles on a sheet of paper. In Part A, the circles are numbers (1, 2, 3, 4 etc.) and the participants are required to connect all the numbers in numerical order. In Part B, the circles are comprised of letters and numbers and the participants must alternate between numbers and letters in sequential order (1, A, 2, B, 3, C, etc.). The time required for the participants to complete each test is used as the score. The Trail Making test is hypothesized to evaluate a variety of cognitive functions including attention, visual search and scanning, sequencing and shifting, psychomotor speed, abstraction, flexibility, ability to execute and modify a plan of action, and ability to maintain two trains of thought simultaneously (Sathouse, 2011; Reger et al., 2004). Trail A and Trail B have been used in previous studies to assess fitness to drive (Austroads., 2004; Classen, Wang, Crizzle, Winter, & Lanford., 2013; Gibbons et al., 2017; Reger et al., 2004). Reliability was found for sub-task A & B respectively (r=0.78 and 0.67) (Goldstein & Watson, 1989) and predictive validity of the on-road driving test was determined for Trail A (p=.002-.05) and Trail B (p=.001-.003) (Marshall et al., 2007).

Clock Drawing Test (CDT). The CDT is a screening tool that evaluates cognitive functions including comprehension, memory, visuospatial abilities, and abstract thinking (Freund, Gravenstein, Ferris, Burke, & Shaheen., 2005). According to Shulman (2000), the CDT yields 85% sensitivity and 85% specificity in identifying cognitive impairment. The task requires participants to draw a clock and set the hands at 10 minutes after 11. The CDT will be scored using a 7-point scoring system (see appendix A). This scoring system has been previously used in assessing fitness to drive (Gibbons et al., 2017; Oswanski et al., 2007). Other studies have used the CDT to evaluate patients driving performance diagnosed with cognitive impairment (Reger et al., 2004).

Modified Mini-Mental State (3MS) Examination. This cognitive screening test offers a brief assessment of an individual's long-term and short-term memory, attention, concentration, orientation to time and place, language ability, constructional praxis, abstract thinking, and list generating fluency (Teng., 1987). The test is comprised of a series of small tests. Other studies have found that the 3MS is a test that measures a patient's general cognition (Reger et al., 2004) Depending on the test taker's level of education, the 3MS yields 95% specificity and sensitivity ranged from 91% - 94% (Teng., 1987). The test is comprised of 15 questions and is scored using a 100-point range.

Experts' clinical decision. The experts' clinical decision was used as the reference standard for the aforementioned cognitive tests. The experts, whose clinical opinions were provided for this study are registered and certified members of The College of Physicians and Surgeons of Ontario, and have numerous years of experience diagnosing and assessing driving abilities for patients with cognitive impairments. The experts have dedicated research time to examine the challenges associated with aging and driving and have presented and published their findings. As practicing physicians working with older adults, they report unfit drivers to the Ministry of Transportation. Their decisions are based on a gestalt approach encompassing a medical examination, cognitive testing, and interviews with the caregiver (i.e., individual accompanying the patient). The medical examination includes a diagnostic dementia evaluation, an evaluation of conditions that preclude to driving (e.g., vision, mobility, seizures, etc.), and in some cases brain imaging. A multitude of cognitive tests are used based on the patient's anticipated areas of difficulty (e.g., MoCA, UFOV, etc.) along with an interview with the caregiver to better understand the patient's driving habits (e.g., accidents, close calls, feeling safe being a passenger, etc.). The expert's clinical decision was recorded as unfit if a copy of the physician's letter

addressed to the Ministry of Transportation was found in the patient's chart indicating their recommendation for driving cessation. Research has demonstrated that expert's predictions on driving cessation has a 90.0 and 60.7 sensitivity and specific respectively, when compared with an on-road driving test (Brown et al., 2005).

Procedure

The retrospective chart review was completed at the Bruyere Memory Program Clinic. Medical charts were pulled and first verified for the patients' signed interest in participating in research. Secondly, charts were reviewed to evaluate if they met the inclusion and exclusion criteria found in Table 1. The CDT was rescored according to Freund CDT Scoring Scale (see appendix A) and the 3MS total score was recalculated to ensure that the scoring was correctly computed. Each qualifying chart proceeded to the next step where the information was recorded. Scoring results from the four neuropsychological tests (i.e., 3MS, Trail A, Trail B, and CDT), demographic data (i.e., age, gender and medical issues that precipitated referral), and the experts' clinical decisions were recorded on a spreadsheet saved on an encrypted USB drive. Charts were examined a second time to verify that they matched the information recorded on the spreadsheet before they were filed to their original location. A final review of the spreadsheet was done ensuring that all participants met the inclusion and exclusion criteria (Table 1) and ineligible participants were removed.

Data Analysis

The data analysis was performed using IBM SPSS Statistic (Version 25; IBM Corp., Armonk, NY). A receiver operating characteristic curve (ROC) using the experts' clinical decision as the reference standard was produced, along with the coordinate points for each test (3MS, Trails A &B, and Clock Drawing) and calculated the area under the curve (AUC). The ROC curve are true

positives (i.e., sensitivity; drivers who are deemed unfit by the neuropsychological test and experts' clinical decision) plotted in function of false positives (i.e., specificity; drivers who are deemed fit by the cognitive test and the experts' clinical opinion) for different cut-off points. Each point on the ROC curve represents a test score with the corresponding level of sensitivity/specify for that particular cut-off point (i.e., coordinate points). However, the AUC does not change depending on a specific cut-off point; rather, it provides an overall summary of the quality of the test validated against the experts' clinical decisions. Using the coordinate points of the ROC, we were able to establish the upper and lower cut-points of each cognitive test that achieved 100% sensitivity and 100% specificity. To establish the upper cut-point that yields 100% sensitivity, the coordinate point on the ROC must equal 1. Conversely, to establish the lower cut-point that yields 100% specificity, the coordinate point on the ROC must be equal to 0.00 (i.e., 100% specificity= 1-specificity). Two cut-off points (i.e., upper and lower) were established for all neuropsychological tests (Trail A, Trail B, 3MS, CDT) individually.

Results

Participant demographics

The sample was comprised of 105 participants who had a mean age of 70.7 years (SD=9.69). Forty-five participants were female (42.9%) and 60 were male (57.1%). Medical issues that precipitated referral to the tertiary memory disorders clinic at Elisabeth Bruyere Hospital were divided into 9 categories including non-specified dementia, MD, DLB, AD, FTD, VaD, MCI, vascular mild cognitive impairment (VMCI), and "Other". Participants who experienced cognitive impairment who were categorized as Other included mild cognitive changes due to a cerebrovascular accident, neurodegeneration, Lyme disease, PD, spread of lymphoma to the central nervous system, and MS. Crosstab frequencies between cognitive impairments and sex

can be found in Table 1. According to the experts' clinical decisions 48.6% (n=51) of participants were evaluated as fit to drive while 51.4% (54) were deemed unfit to drive.

Table 2
Cognitive Impairments in Female and Male Participants.

Cognitive Impairments	Female	Male	Total
Non specified Dementia	7	9	16
Mixed Dementia	3	1	4
MCI	11	12	23
LBD	1	4	5
AD	6	8	14
VMCI	10	19	29
FTD	2	2	3
VaD	1	0	1
Other	4	5	10
Total	45	60	105

Note. MCI= Mild Cognitive Impairment; DLB= Dementia with Lewy Bodies; AD= Alzheimer's disease; VMCI= Vascular Mild Cognitive Impairment; FTD= Frontotemporal Dementia; VaD= Vascular Dementia.

Neuropsychological Tests and Cut-Off Scores

During the retrospective chart review, data were collected on participants' test scores including the CDT, the 3MS, and the Trail Making Test Part A and B. Descriptive statistics of the neuropsychological test can be found in Table 2. Furthermore, cut-off points were established to yield 100% sensitivity, specificity, positive predictive value, and negative predictive values. For each test, coordinate points of the ROC curve were produced using experts' clinical decision as a reference standard. Cut-off points that reflected 100% specificity were used to categorize patients who would have been evaluated as unfit to drive according to the experts' clinical decision, while cut-off points that reflected 100% sensitivity were used to categorize patients who were fit to drive. The percentage of indeterminates represent the number of participants who did not obtain

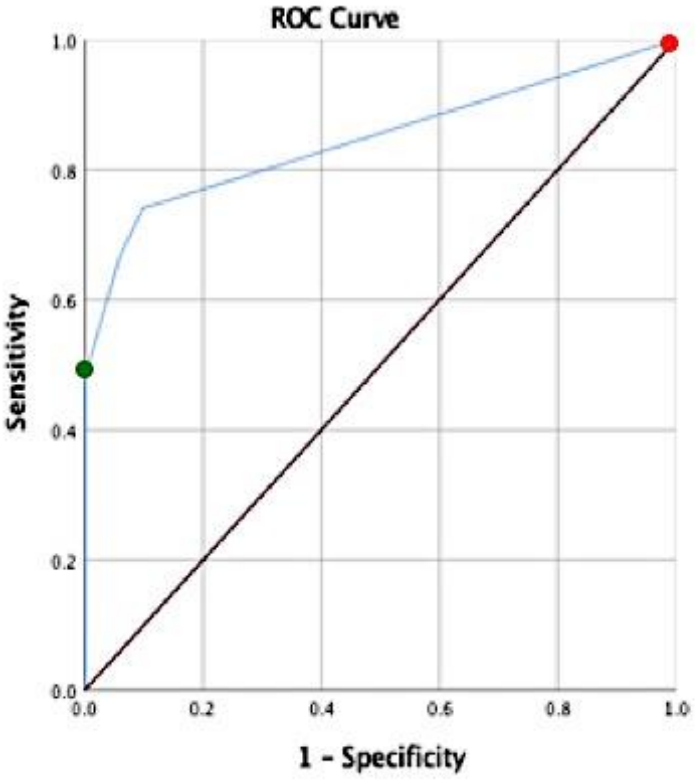
sufficient test scores to be evaluated as fit to drive, while neither were their scores low enough for them to be evaluated as unfit to drive.

Table 3
Neuropsychological Tests' Cut-off points.

Test	M	SD	Min.	Max.	100% Sensitivity[a]	100% specificity[b]	% Indeterminate[c]	AUC [95% CI]
CDT	5.44	2.30	0	7	7	4.5	84%	.846[.769, .924]
3MS	84.67	14.21	21	99	97.5	77.5	84%	.867[.800, .934]
Trail B[d]	186.71	87.43	43	300	79.5	293.5	75.6%	.942[.898, .987]
Trail A[d]	76.31	63.11	24	300	38.5	90.5	70.4%	.884[.823, .946]

Note. M=mean; SD= standard deviation; Min.= minimum value; Max.= maximum value; AUC= area under the curve; CDT= Clock drawing test; 3MS= Modified Mini-Mental State Examination; Trail B= Trail Making Test Part B; Trail A= Trail Making Test Part A. [a]Cut-off scores indicate unfit drivers on the basis of the expert's clinical decisions, participants obtaining test scores equal or greater than this value for Trail B and Trail A and equal or lower than this value for CDT and 3MS. [b]Cut-off scores indicate fit drivers on the basis of the expert's clinical decisions, participants obtaining test scores equal or lower than this value for Trail B and Trail A and equal or greater than this value for CDT and 3MS. [c]Percentage of participants who cannot be rated as fit or unfit drivers by the test demonstrating 100% sensitivity and 100% specificity on the basis of expects clinical decisions. [d]Measured in seconds.

Clock Drawing Test. The ROC curve for the Clock Drawing Test along with the curve's coordinate points can be found in Figure 1. The area under the curve was AUC= .846 (95% CI, .769, .924). Coordinate points of the curve were used to establish 100% sensitivity cut-off points and 100% specificity cut-off points. Based on the Freund CDT scoring scale, cut-off points were > 8 and < 4.5, respectively.

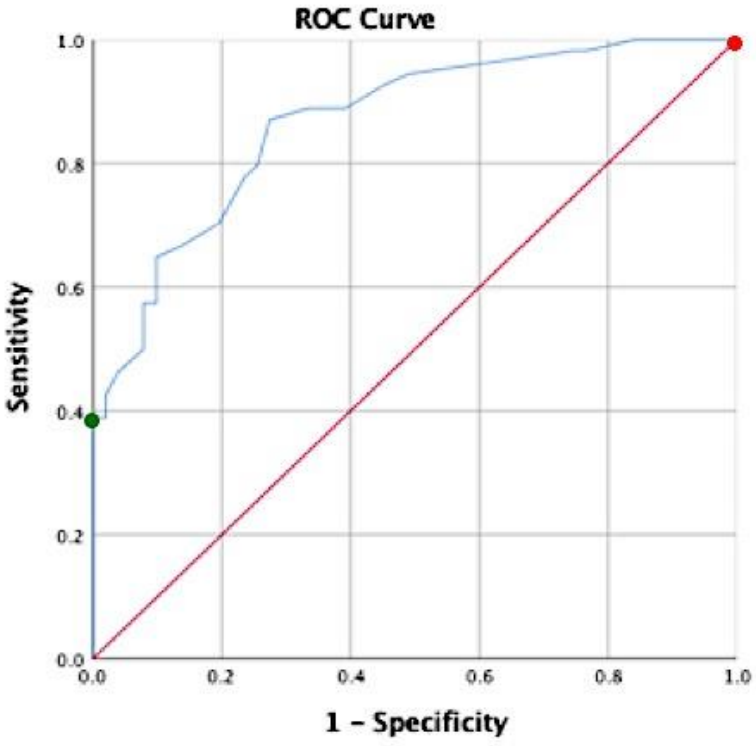


Coordinate Points

Test scores	Sensitivity	1-Specificity
-1.00	.000	.000
.50	.185	.000
1.50	.241	.000
2.50	.259	.000
3.50	.315	.000
4.50	.481	.000
5.50	.667	.059
6.50	.741	.098
8.00	1.00	1.00

Figure 2.
Receiver operating characteristic (ROC) curve with coordinate points for the Clock Drawing Test (CDT).
Note. The red circled coordinate points reflect the upper cut-off point indicated in red on the ROC curve that yields 100% sensitivity. The green circled coordinate points reflect the lower cut-off point indicated in green on the ROC cure that yields 100% specificity. The blue line on the ROC curve represents indeterminate test scores. The diagonal line symbolizes 0% sensitivity and specificity.

Modified Mini-Mental State (3MS) Examination. The ROC curve for the 3MS along with the curve's coordinate points can be found in Figure 2. The area under the curve was AUC=.867 (95% CI, .800, .934). Coordinate points of the curve were used to establish 100% sensitivity and 100% specificity cut-off points. Based on the 3MS scoring system, cut-off points were >97.5 and <77.5, respectively.



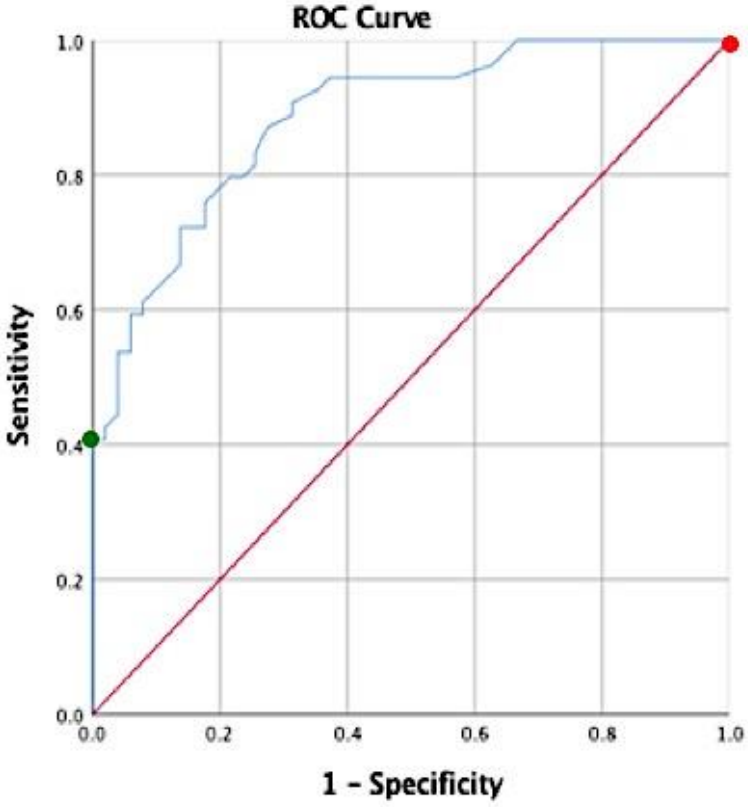
Coordinate Points

Test Scores	Sensitivity	1-Specificity	Test Scores	Sensitivity	1-Specificity
-1.00	.000	.000	80.50	.463	.039
10.50	.019	.000	81.50	.500	.078
28.50	.037	.000	82.50	.574	.078
38.50	.056	.000	83.50	.574	.098
46.00	.074	.000	84.50	.593	.098
52.00	.093	.000	85.50	.648	.098
54.00	.111	.000	86.50	.667	.137
55.50	.130	.000	87.50	.704	.196
57.50	.148	.000	88.50	.778	.235
62.50	.185	.000	89.50	.796	.255
67.00	.204	.000	90.50	.870	.275
69.50	.241	.000	91.50	.889	.333
71.50	.259	.000	92.50	.889	.392
72.50	.278	.000	93.50	.926	.451
73.50	.296	.000	94.50	.944	.490
74.50	.315	.000	95.50	.981	.745
75.50	.352	.000	96.50	.981	.765
76.50	.370	.000	97.50	1.000	.843
77.50	.389	.000	98.50	1.000	.922
78.50	.389	.020	100.0	1.000	1.000
79.50	.426	.020			

Figure 3. Receiver operating characteristic (ROC) curves with coordinate points for the Modified Mini-Mental State (3MS) Examination.

Note. The red circled coordinate points reflect the upper cut-off point indicated in red on the ROC curve that yields 100% sensitivity. The green circled coordinate points reflect the lower cut-off point indicated in green on the ROC cure that yields 100% specificity. The blue line on the ROC curve represents indeterminate test scores. The diagonal line symbolizes 0 % sensitivity and specificity.

Trail Making Test Part A. The ROC curve for the Trail Making Test Part A along with the curves coordinate points can be found in Figure 3. The area under the curve was AUC=.884 (95% CI, .823, .946). Coordinate points of the curve were used to establish 100% sensitivity and 100% specificity cut-off points. Based on the time (in seconds) to complete the Trail Making Test, cut-off points were <38.5 and >90.5, respectively.



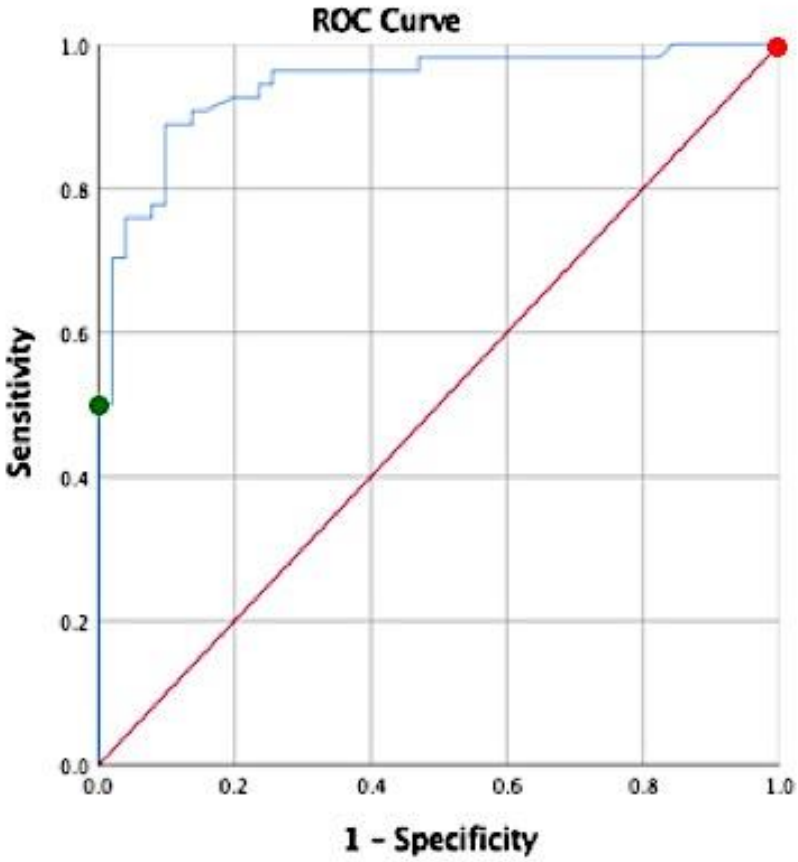
Coordinate Points

Test Scores	Sensitivity	1-Specificity	Test Scores	Sensitivity	1-Specificity
23.00	1.000	1.000	64.50	.648	.118
25.50	1.000	.961	65.50	.611	.078
27.50	1.000	.922	66.50	.593	.078
28.50	1.000	.902	68.50	.593	.059
29.50	1.000	.882	71.50	.574	.059
31.50	1.000	.863	74.50	.556	.059
34.00	1.000	.843	76.50	.537	.059
35.50	1.000	.784	77.50	.537	.039
36.50	1.000	.725	79.50	.481	.039
37.50	1.000	.686	80.50	.444	.039
38.50	1.000	.667	81.50	.426	.039
39.50	.963	.627	82.50	.407	.020
40.50	.944	.569	86.50	.389	.020
41.50	.944	.510	90.50	.389	.000
42.50	.944	.412	97.00	.333	.000
43.50	.944	.373	103.50	.315	.000
44.50	.926	.353	106.00	.296	.000
46.00	.907	.314	114.00	.278	.000
47.50	.889	.314	123.00	.259	.000
48.50	.870	.275	130.50	.241	.000
50.00	.833	.255	137.00	.222	.000
51.50	.815	.255	139.50	.204	.000
52.50	.796	.235	143.50	.185	.000
53.50	.796	.216	159.00	.167	.000
55.50	.778	.196	175.50	.148	.000
57.50	.759	.176	184.00	.130	.000
58.50	.722	.176	224.00	.111	.000
59.50	.722	.137	280.00	.093	.000
61.50	.685	.137	301.00	.000	.000
63.50	.667	.137			

Figure 4. Receiver operating characteristic (ROC) curves with coordinate points for the Trail Making Test Part A.

Note. The red circled coordinate points reflect the lower cut-off point indicated in red on the ROC curve that yields 100% sensitivity. The green circled coordinate points reflect the upper cut-off point indicated in green on the ROC curve that yields 100% specificity. The blue line on the ROC curve represents indeterminate test scores. The diagonal line symbolizes 0 % sensitivity and specificity.

Trail Making Test Part B. The ROC curve for the Trail Making Test Part B Test along with the curve's coordinate points can be found in Figure 3. The area under the curve was AUC=.942 (95% CI, .898, .987). Coordinate points of the curve were used to establish 100% sensitivity and 100% specificity cut-off points. Based on the time (in seconds) to complete the Trail Making Test, cut-off points were <79.5 and >293.5, respectively.



Coordinate Points

Test scores	Sensitivity	1-Specificity	Test scores	Sensitivity	1-Specificity
42.00	1.000	1.000	147.50	.926	.235
51.50	1.000	.980	149.50	.926	.216
61.00	1.000	.941	153.00	.926	.196
64.50	1.000	.922	155.50	.907	.157
69.50	1.000	.902	156.50	.907	.137
74.50	1.000	.882	158.50	.889	.137
78.00	1.000	.863	162.00	.889	.098
79.50	1.000	.843	165.00	.870	.098
82.50	.981	.824	167.50	.852	.098
85.50	.981	.784	172.50	.833	.098
86.50	.981	.745	176.50	.815	.098
87.50	.981	.706	182.50	.796	.098
89.00	.981	.667	189.50	.778	.098
91.00	.981	.647	193.00	.778	.078
93.00	.981	.627	196.50	.759	.078
95.00	.981	.569	201.50	.759	.039
96.50	.981	.510	207.50	.741	.039
98.00	.981	.490	220.00	.722	.039
99.50	.981	.471	230.50	.704	.039
101.50	.963	.471	234.00	.704	.020
104.00	.963	.451	237.50	.685	.020
106.00	.963	.431	240.50	.667	.020
107.50	.963	.412	247.50	.648	.020
109.00	.963	.392	254.00	.630	.020
111.50	.963	.373	258.50	.611	.020
114.00	.963	.353	261.50	.593	.020
123.00	.963	.333	267.00	.574	.020
133.00	.963	.314	272.50	.556	.020
135.50	.963	.294	275.50	.537	.020
138.00	.963	.275	281.00	.519	.020
141.50	.963	.255	285.50	.500	.020
143.50	.944	.255	293.50	.500	.000
145.50	.944	.235	301.00	.000	.000

Figure 5. Receiver operating characteristic (ROC) curves with coordinate points for the Trail Making Test Part B.

Note. The red circled coordinate points reflect the lower cut-off point indicated in red on the ROC curve that yields 100% sensitivity. The green circled coordinate points reflect the upper cut-off point indicated in green on the ROC curve that yields 100% specificity. The blue line on the ROC curve represents indeterminate test scores. The diagonal line symbolizes 0 % sensitivity and specificity.

Discussion

In Canada, car dependency is substantial as it is a sparsely populated country often requiring individuals to travel by private vehicle to meet their needs (Statistics Canada, 2007). Driving and having access to a private vehicle have many benefits; however, this also comes at a cost. While the car increases accessibility, flexibility, autonomy, control, and longevity (Ellaway et al., 2013; Smith et al., 1990) it also contributes to a high number of fatalities and personal injuries (Statistics Canada, 2016). Adults over the age or 65 are at a greater risk of being involved in a collision and as the aging population continues to grow, the proportion of older drivers will increase (Statistics Canada, 2012). Age alone does not negatively impact driving abilities but rather secondary effects related to aging can have a negative impact on driving ability. Certain conditions, notably, cognitive impairment become more prevalent with age and reduce one's ability to drive safely (Tuokko & Hunter, 2002). Dementia is the leading cause of cognitive impairment among older adults and effects the patient's memory, attention, visuospatial skills, and executive function.

The aforementioned cognitive abilities are crucial in order to engage in safe driving. It is essential to find a balance between road safety and mobility amongst older adults since certain individuals are capable of driving safely despite cognitive impairment. Primary care physicians hold a legal responsibility of making that decision. However, most general practitioners reported not being confident in assessing fitness to drive and the majority expressed they could benefit from further education and better guidelines to support their clinical judgement (Dickerson et al., 2017; Herman, 2006; Jang et al., 2007). This rising concern has motivated researchers to develop a valid and reliable off-road driving test that could assist physicians in making decisions around fitness to drive. Over the years, many attempts have been made without much success. Tests often lacked sensitivity or specificity, thus, misclassifying drivers as either fit or unfit to drive. The inability of

establishing appropriate cut-off points on cognitive tests has been the main culprit for insensitive and unspecific tests. These mistakes have large consequences affecting public safety or reducing the quality of life of patients. There has been great interest in developing a reliable and valid off-road test that could optimize efficiency, provide a safer way to assess fitness to drive (Gibbons et al., 2017) and facilitate routine assessments (Unsworth et al., 2018).

Serial trichotomization has demonstrated potential in being an excellent tool for physicians that could optimize the assessment of fitness to drive, in that it yields 100% sensitivity and specificity. However, further research is required as serial trichotomization has been the subject of limited research. Multiple studies are required to confirm validity and reliability of an assessment method and replication is strengthened if clinical setting and sample size differ. This study aims to add to the growing body of literature supporting the use of serial trichotomization by comparing the findings using a different clinical setting, enlarged sample size, and an experts' clinical decisions as a reference standard. Under the conditions stated above, we hypothesized that this study would draw similar cut-off points to the findings of Gibbons and colleagues (2017) and would have the potential of categorizing a comparable percentage of drivers as fit or unfit to drive without having to undergo further testing.

Neuropsychological Tests and Cut-off Points

In line with our hypothesis, the current study drew similar cut-off scores to the previous study having trichotomized the scores of common neuropsychological tests (Gibbons et al., 2017) set to yield 100% sensitivity and specificity. In both studies, similar cut-off points were established for the CDT, both having the exact same upper cut-off point (>7). The upper and lower cut-off points for the 3MS were established at >98 and <78 respectively. In the previous study, the Montreal Cognitive Assessment (MoCA) was used instead of the 3MS in our study. However,

according to the crosswalk scores from the MOCA to the 3MS (Sink et al., 2015), both cut-off points (i.e., upper and lower cut-off points) were the same. A crosswalk enables the comparison of results between two different tests by providing an equivalent value for every score. In the Trail Making Test Part B, the upper cut-off points were the same (< 80s) in both studies, however, the lower cut-off points were quite different when comparing the current study (>293s) with the previous study (>178s). Cut-off points for the Trail Making Test Part A were different for the upper and lower cut-off points. It is important to keep in mind that the Trail Making Test Part A and Part B are scored in seconds, thus, the differences between cut-off points in both studies can appear much larger than they actually represent. A more extensive comparison of cut-off points can be found in Table 4.

Table 4
Comparison of Cut-Off Points to Previous Study Having Trichotomized Common Neuropsychological Test Scores.

	Sensitivity Cut-Off Points		Specificity Cut-Off Points	
Cognitive Tests	Current Study	Previous Study [a]	Current Study	Previous Study [a]
CDT	7	7	5	4
3MS	98	97-98 [b]	78	77-79 [b]
Trail B [c]	80	80	293	178
Trail A [c]	38	25	90	69

Note. CDT= Clock drawing test; 3MS= Modified Mini-Mental State Examination; Trail B= Trail Making Test Part B; Trail A= Trail Making Test Part A. [a] Study conducted by Gibbons et al. (2017). [b] Cut-off points originally scored on the Montreal Cognitive Assessment (MOCA) transferred to the 3MS using the crosswalk from Sink et al. (2015). [C]Measured in seconds.

There are at least two potential explanations that can account for the differences in the cut-off points between both studies. Looking at the general trending differences, cut-off points were established at more conservative levels (i.e., higher upper and lower cut-off scores) than in the previous study, with the exception of the CDT. The spectrum bias could offer a potential explanation for the higher upper and lower cut-off points found in the study conducted by Gibbons and colleagues (2017). The spectrum bias refers to the phenomenon where the performance of a classification test (i.e., sensitivity and specificity) varies according to the sample of participants used in the study and their clinical profiles. Sensitivity and specificity are not affected by the changes in prevalence (i.e., number of fit or unfit drivers), but are affected by the changes in the spectrum of participants' cognitive abilities that inherently have implications on the decision regarding fitness to drive. The sensitivity and specificity of a test, when it is used to differentiate a sample of participants who obviously do have cognitive impairments to participants who obviously do not have cognitive impairment, likely overestimate its performance in comparison to a test that is applied in a clinical context where there is a spectrum of cognitive impairment (i.e., mild, moderate, & severe) (Weaver et al., 2013; Usher-Smith, Sharp, & Griffin, 2016). Within the realms of the serial trichotomization method, sensitivity and specificity are viewed as independent as they must always yield 100%, thus, making cut-off scores dependent. To conclude, in studies using more clear-cut clinical profiles, cut-off points are likely overly conservative.

In the previous study using the serial trichotomization method (Gibbons et al., 2017), the sample of participants was comprised of patients with a heterogeneous clinical profile. Some participants did not have any obvious clinical profile of cognitive impairment and the reason for their referral was based on physical injury, chronic pain, spinal injury, or amputation of leg or arm, while others had a more obvious clinical profile of cognitive impairment (e.g., cerebrovascular

accident, traumatic brain injury and cognitive deficits). In contrast, participants in our sample were all diagnosed with a form of cognitive impairment that varied on a spectrum of severity (e.g., MCI, mild dementia, moderate dementia, severe dementia). It is plausible that the spectrum bias could partially account for the difference in higher cut-points between both studies. It has been argued by many researches that off-road driving tests should be designed to test specific medical conditions and therefore be validated using a sample that reflects that same condition (Dickerson et al., 2017). This study adds to the growing body of literature by potentially facilitating the establishment of more precise cut-off points specific to patients who may experience cognitive impairment. Research has demonstrated that older adults with moderate and severe dementia should not drive and are at greater risk of being involved in a collision. Assessing fitness to drive in those situations has been proven to be easier and more clear-cut. However, evaluating patients' driving abilities with early and mild dementia poses a dilemma (Dickenson, 2014). Establishing more precise cut-off points on cognitive tests can help provide a better assessment regarding fitness to drive where it is most needed.

Acknowledging that both studies have used different reference standards to determine validity is another plausible explanation that could have accounted for the differentiation among cut-off points. The current study used experts' clinical decisions as the criterion to establish cut-off points while the previous study used the gold standard of an on-road driving test. It is probable that the different reference standards could have partially accounted for the differences in cut-off points. Despite the aforementioned difference, the AUC of each cognitive test for both studies yielded similar results. The AUC provides an overall summary of the quality of the test validated against the reference standard (e.g., on-road driving test, simulator, experts' clinical decisions). According to the interpretation guidelines of AUC values proposed by Hosmer and Lemesshow

(2000) (Weaver et al., 2013), the CDT, 3MS or MOCA were classified as excellent tests (i.e., AUC < .8) and the Trail Making Test Part B was classified as an outstanding test (i.e., AUC< .9) in determining fitness to drive based on experts' clinical decisions and on-road driving tests (Gibbons et al., 2017). However, the Trail Making Test Part A was classified as an outstanding test when validated against the on-road driving test in comparison to an excellent test when used against the experts' clinical decision.

If common neuropsychological tests are to be used to assess fitness to drive, they must yield valid and reliable cut-off points, thus consistently categorizing drivers as fit or unfit based on their ability to operate a motor vehicle. As mentioned previously, tools that are used to inform fitness to drive need to be validated against a reference standard. Unfortunately, no reference standard is perfect and despite establishing cut-off points that yield 100% sensitivity and specificity, misclassification errors are still possible due to imperfections of the reference standard (Kay et al., 2012). This can create a bias in the performance of a classification test (Weaver, 2013). Comparing cut-off scores established by different reference standards can be used to dissipate some bias and increase reliability as each reference standard has their own advantages and disadvantages. In the current study, cut-off points were established against experts' clinical decisions, however in a previous study, the cut-off points were established using the gold standard of an on-road driving test (Gibbons et al., 2017). On-road driving test are not standardized as testing conditions are near impossible to control (e.g., maneuvers are dependent on area, traffic flow, impact of other drivers) (Kay et al., 2012; Reger et al., 2004; Shechtman et al., 2010). Clinical experts also do not have a standardized or widely used protocol and decisions are made in a gestalt manner based on multiple evaluations (Brown et al., 2005). Despite those ambiguities, both studies have yielded similar cut-off points demonstrating reliability in the cut-off points established for

the CDT, 3MS or MoCA and Trail B (upper cut-off point) and limited bias provided from the reference standard affecting the levels of sensitivity and specificity. Other studies have also combined information from multiple imperfect reference standards for validation of screening instruments (Pence, Miller, & Gaynes, 2010).

Observing each neuropsychological test independently and their cut-off points has provided some interesting insight into the efficacy of certain tests. The CDT was unable to classify 85% of drivers as fit or unfit. Despite having established the same cut-off scores and using the same scoring system (i.e., Freund CDT Scoring Scale), the study conducted by Gibbons and colleagues (2017) were unable to classify 63.4% of participants regarding fitness to drive. It is highly probable that the different clinical profile of participants can account for the variation in the percentage of indeterminates. Manning, Davis, Papandonatos, & Ott (2014), have found similar results using the CDT as a screening measure of driving performance in 122 healthy and cognitively impaired older drivers. Using the Freund's 7-point scoring system, the CDT was unable to classify 76% of participants with an upper cut-off point (>6) and a lower cut-off point (< 3). Other studies have shown comparable AUC =.90 (95% CI, .82, .95) when using the errors measured by driving simulator as a reference standard (Freund et al., 2005) while others have found relatively low AUC=.62 (95% CI, .51, .71) when comparing against an on-road driving test (Manning et al., 2014). It is still debatable whether the CDT is a sufficient test to evaluate fitness to drive in patients with cognitive impairment. In the realm of the current study and of a previous study (Gibbons et al., 2017), the CDT was only able to discriminate between drivers who were unfit to drive using the lower cut-off point (< 5 or <4), however, the CDT was unable to identify drivers who were fit to drive with 100% sensitivity as the upper cut-off point (>7) was impossible to obtain since the maximum score on the scoring scale was 7.

Similar results were found for the Trail Making Test Part A and B. In the current study, the Trail A was unsuccessful at making 70.4% of decisions regarding fitness to drive and Trail B failed to categorize 75.6% as ether fit or unfit. Gibbons and colleagues (2017) yielded better results with a more heterogeneous sample of participants. Trail A was unable to categorize 66.3% of participants, while Trail B demonstrated the best results with 33.7% of participants that required further testing. Dobbs and Shergill (2013) attempted to evaluate the efficacy of the Trail Making Test (Part A and B) in identifying drivers with cognitive impairment. The study included 87 healthy controls and 47 cognitively impaired patients, and found that the Trail Making Test power is lowest when identifying unfit drivers with cognitive impairment. This could partially explain our higher levels of indeterminates as our sample was only comprised of patients with cognitive impairment contrary to the previous study (Gibbons et al.,2017).

The Modified Mini-Mental State (3MS) Examination has not been used as a tool to determine fitness to drive, however, the Mini-Mental State Examination (MMSE) has been extensively used as a tool to assess fitness to drive (Hollis et al., 2015; Joseph et al., 2014; Badenes et al., 2018). Studies suggest that the 3MS showed improved classification of cognitive impairment accuracies relative to the MMSE (Patten, Britton, & Tremont, 2018). According to the AUC= .867 (95% CI, .800, .934), the 3MS is classified as an excellent test in assessing fitness to drive using experts' clinical decision as a reference standard. However, the 3MS was only able to make 16% of decisions regarding fitness to drive. In the previous study that had utilized the serial trichotomization model, they had replaced the 3MS with the MoCA and found slightly better results. The MoCA was only able to inform 20% of decisions related to fitness to drive (Gibbons et al., 2017).

The findings in this study add to the existing literature that have determined that one cognitive test is not sufficient to make a decision regarding fitness to drive (Bédard & Dickerson, 2014; Davis et al., 2016; Manning, 2014). Trichotomization of cognitive tests have shown to establish better cut-off points with fewer misclassification errors, whereas using only one test results in a large grey area, thus obligating the majority of patients to undergo further testing in order to make a determination on their ability to drive safely. In the next section, we further discuss the efficacy and implications of trichotomizing multiple subsequent tests together to replicate a similar serial trichotomization model that was previously tested (Gibbons et al., 2017).

Funnelling Model

The funnelling model was constructed using the cut-off points established on the common neuropsychological tests above. Cut-off points were rounded-off to yield possible scores on the neuropsychological test in order to facilitate interpretation and classification. Following the trichotomization method, participants' test scores were categorized into three groups: pass, fail, and indeterminate. Participants whose test scores were superior to the upper cut-point were categorized in the passing group and participants whose test scores were inferior to the lower cut-point were categorized in the fail group. Finally, participants whose test scores were inferior to the upper cut-point but superior to the lower cut-point were categorized as indeterminate. This categorization was done individually for all cognitive tests. Cognitive tests were then placed in ascending order according to their percentage of indeterminate participants (i.e., Trail A, Trail B, 3MS, and CDT). In other words, the test with the lowest percentage of indeterminate participants was placed first finishing with the test with the highest percentage of indeterminate participants. This was only done for the purpose of determining the order of tests to form the serial trichotomization model. This first step was completely independent of the next step.

In the model (see figure 6), participants were funneled through the first cognitive test, Trail A, with the smallest grey area (i.e., smallest percentage of indeterminate participants). Participants whose test scores for the first test were < 38s passed and were deemed fit to drive. On the contrary, those whose scores were > 90s failed and were classified as unsafe to drive. The participants whose scores on the Trail A test fell in the grey area (i.e., between both cut-off points) were categorized as indeterminate and were funneled through the next test; Trail B. On the basis of the Trail A test, 20% (n=21) of participants were evaluated as fit to drive and 16.2% (n=17) were classified as unfit to drive. Sixty-three percent (n=67) of the participants could not be classified using the Trail A test with 100% sensitivity and specificity and proceeded to the next level of the serial trichotomization model. The same steps were taken for the Trail B test where nine participants earned a score >293 and were categorized as unfit to drive, while five participants scored < 80 and were deemed fit to drive. Fifty-three percent of participants were classified as indeterminate and moved on to the next test. Four participants obtained scores <78 and where deemed unfit to drive on the 3MS while five participants yielded scores >98 and were categorized as fit to drive. Forty-two percent of participants (n=44) remained in the indeterminate category at this level of the model. The aforementioned participants were funnelled through the last test included in the model. The CDT categorized six participants as unfit to drive and no participants as fit to drive. After participants were funnelled through all levels of the model, 38.1% (n=40) of participants were evaluated as unfit to drive while 25.8% of participants (n=27) were deemed fit to drive. Only 36.1% of participants (n=38) remained indeterminate and required further testing.

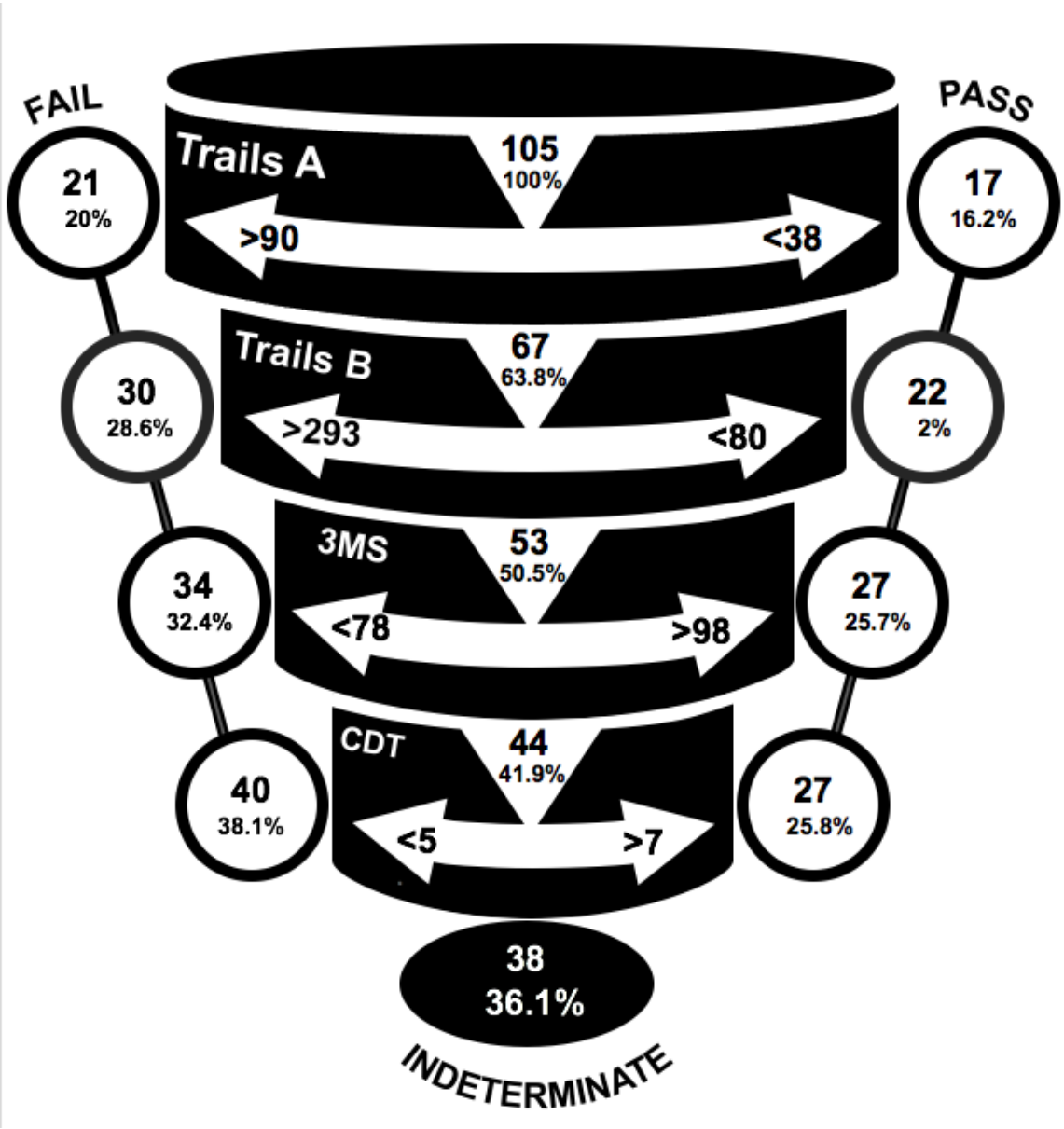


Figure *6.* Funnelling model used for the serial trichotomization method proposed by Gibbons and colleagues (2017). Participants' test scores which do not correspond to the cut-off scores of pass or fail are categorized as indeterminate scores. Indeterminate scores are filtered through the next tests. Adapted from "Using Serial Trichotomization with Common Cognitive Tests to Screen for Fitness to Drive", by C. Gibbons, N. Smith, R. Middleton, J. Clack, B. Weaver, S. Dubois, and M. Bédard, 2017, *American Journal of Occupational Therapy*, 71(2). Copyright 2017 by the American Psychological Association.

Note. CDT= Clock Drawing Test; 3MS= Modified Mini-Mental State Examination; Trail B= Trail Making Test Part B; Trail A= Trail Making Test Part A.

In line with our second hypothesis, we were able to classify a comparable number of participants in terms of fitness to drive using a similar serial trichotomization model in comparison to the previous study (Gibbons et al., 2017). The model managed to make 63.9 % (n=67) of the decisions around participant's abilities to drive safely without needing to undergo further testing with 100% sensitivity, specificity, positive predictive value, and negative predictive values. In the previous study conducted by Gibbons and colleagues (2017), a similar model managed to categorize 78.3% (n=65) of participants as either fit or unfit to drive. It is important to note that the previous serial trichotomization model was comprised of 5 common neuropsychological tests (i.e., Trails B, CDT, MVPT-3, Trails A, MoCA). Conversely, the current model only included four common neuropsychological tests (Trails B, CDT, Trails A, 3MS). Each neuropsychological test reduced the number of indeterminate participants. Thus, it is probable that the current study would have been able to yield better results if the model included more neuropsychological tests. Furthermore, this claim can be supported by other studies that have used single trichotomization models to assess the fitness to drive of patients with cognitive impairment. In the first study, the DriveAware test was only able to make 50% of decisions regarding older adults' fitness to drive (Kay et al., 2009). Similarly, the SIMARD test only managed to classify 49% of participants as either fit or unfit to drive (Dobbs & Schopflocher, 2010). This supports the claim that the use of one single test is not sufficient in predicting fitness to drive as it results in a large grey area (i.e., participants require further testing) and every subsequent test added to the model helps to mitigate the proportion of grey area, increasing the number of decisions made regarding fitness to drive (Bédard & Dickerson, 2014).

Other factors could have also contributed to the more elevated number of indeterminate classifications. As mentioned previously, participants in the sample had different clinical profiles thus influencing the cut-off points and proportion of grey area (i.e., percentage of indeterminate) for some of the neuropsychological tests. Inherently, this could also have had an impact on the number of participants that were classified as indeterminate since the grey area of the Trail Making Test Part A and B were larger than in the previous study (Gibbons et al., 2017). Furthermore, a bigger sample size was used in the current study (n=105) in comparison to the previous study (n=83). Using a bigger sample size yields more representative results and has the potential of establishing more precise cut-off points as it increases the chances of including a bigger spectrum of impairment. It is plausible that the sample size could have partially accounted for the difference in the percentage of participants categorized as indeterminate among the two studies.

The serial trichotomization method has demonstrated the ability to classify older drivers with cognitive impairment as either fit or unfit to drive yielding great performance measurements (i.e, 100% sensitivity and 100% specificity). Many researchers support that a single screening tool is not sufficient to make a decision regarding fitness to drive, mainly because driving involves a combination of multiple skills (e.g., attention, visual scanning, reaction time, etc.) and a single test may not evaluate all abilities required for driving (Bédard & Dickerson, 2014). Referring to the current serial trichotomization model, patients' driving abilities may only be evaluated as fit to drive by one test. For example, if the patient scores < 38s on the Trail A test, they are automatically deemed fit to drive and are excluded from all other subsequent tests. This could be problematic as every individual test is involved in testing specific abilities that are related to driving performance. The Trail A test mainly evaluates cognitive domains associated with attention and concentration, but does not evaluate other cognitive domains that are critical in driving safely (e.g., visuospatial

skills, and memory) (Reger et al., 2014). In some neurodegenerative diseases (e.g., mild VaD), precise cognitive domain (e.g., executive functioning) can be significantly impaired without having a significant effect on others (e.g., memory, attention) (Korczn et al., 2012; Nordlund et al., 2007). A patient with a similar profile could have difficulty with set-shifting abilities that are crucial in safe driving and still score sufficiently well on the Trail A test and pass, therefore being exempt from all other tests.

We propose a modification to the serial trichotomization model (Gibbons et al., 2017) to avoid making decisions regarding fitness to drive solely based on one test. The model should include a disclaimer indicating that patients who are categorized as pass must yield superior scores to the lower-cut off point on all neuropsychological tests included in the model. Older adults who pass one test who later yield a score sufficiently low to fail another test, should be deemed unfit to drive. This would assure that decisions surrounding fitness to drive would not be done based on only one neuropsychological test. In the current study, no participants that have passed one test later failed another. However, in both studies that utilized the serial trichotomization method, the evaluators were not blinded to the results of the cognitive tests, thus, potentially mitigating the occurrence of the aforementioned situation. (A more detailed explanation of the impact of this bias can be found in the limits section.) The disclaimer should be included in forthcoming studies utilizing the serial trichotomization method to inform questions surrounding fitness to drive.

Researchers have found that most primary care physicians who address the subject of driving with their patients suffering from cognitive impairment refer them to a driving evaluation service (Adler & Rottunda, 2011). A complete assessment is costly and can range from five hundred dollars to eight hundred dollars and is not subsidised by government health care programs in Canada. Individuals may also be referred to a Medical Advisory Committee where a panel of

expert medical advisors review complex cases and provide an expert opinion (Ministry of Transportation). The current assessment procedures result in long waiting times and lengthy testing periods (Kay et al., 2012). Many research efforts have been made in recent years to address the difficulties associated with assessing medically at-risk older drivers by primary care physicians.

The Canadian Driving Research Initiative for Vehicular Safety in the Elderly (Candrive), is a research program that was established in 2002 with the aim to develop, implement and validate a screening tool that is easy to use and allows primary care physicians to make better decisions around fitness to drive for older adults. The 6-year longitudinal study aims to include pencil-and-paper tests that are easy to administer and score (Marshall et al., 2013). The serial trichotomization method has recently demonstrated great potential in practical application by assisting primary care physicians in making decisions around fitness to drive among older adults with cognitive impairments (Gibbons et al., 2017). This model could offer physicians a better preliminary method in assessing fitness to drive using common paper-pencil cognitive tests. With easy administration, a simple scoring system, and no equipment requirement, serial trichotomization allows primary care physicians to make specific and sensitive decisions around fitness to drive, in a timely fashion. The current study adds to the growing body of literature supporting the use of serial trichotomization by yielding similar results on cut-off points and classifying a comparable percentage of drivers as either fit or unfit without having to undergo further testing. Further research is needed before the serial trichotomization method can be implemented to aid primary care physicians to make decisions surrounding fitness to drive in older adults.

Limitation and Future Research

There are several limitations surrounding this study that should be considered when reviewing the results. To begin, the retrospective nature of the research design is a limitation. Since

the data extracted from the medical charts were not recorded for research purposes, comprehensiveness and quality were at times deficient requiring the following charts to be excluded from the study. This could have inherently contributed to a selection bias as the selected sample may have not entirely been representative of the population intended to be studied. Medical charts needed to be individually selected to meet inclusion and exclusion criteria, and comprehensiveness needed to be ensured before abstraction of the data began. Furthermore, the abstractor was not blinded to the study's objectives and hypothesis, thus, introducing an abstractor bias. In situations where the abstractor is not blinded to the study, they may be more drawn to certain data while overlooking less desirable data. Efforts were made to mitigate this bias by creating a systematic review and establishing strict inclusion and exclusion criteria before data collection. Also, the data collection was performed with only one abstractor making it difficult to quantify consistency in the data collection. To help eliminate the number of possible errors while extracting data, at least two abstractors should independently record the data in order to ensure accuracy. Given the vast complications associated with a retrospective research design, future research should concentrate on adopting a prospective research design as they are generally considered to yield the best research quality (Abbott, Barton, Terhorst, & Shembel, 2016; Hess, 2014; Kaji, Schriger, & Green, 2014).

Another important limitation to consider is the reference standard used in the current study. The experts' clinical decisions were used as the reference standard to inform decisions related to fitness to drive. The experts were not blinded to the results of the patients' cognitive scores when determining their fitness to drive, thus introducing an incorporation bias. It is preferable that the reference standard and the screening test assessed are completely independent of one another. When there is a lack of independence between both measures there is an overestimation of the

accuracy of the screening test. This also translates to the overestimation of cut-off points. The incorporation bias has a particularly large impact on the results of this study since the test characteristics are excellent. Little impact is accounted for when test characteristics are less favorable since they would have been even worse than one would expect if the incorporation bias was not present. However, implementation of an excellent test including an incorporation bias could be costly, leading to misclassification errors. Both studies that have evaluated the use of serial trichotomization were subject to an incorporation bias. Future research is required using a model where those who determine fitness to drive are blinded to the neuropsychological test scores to mitigate the influence of an incorporation bias (Gibbons et al., 2017).

The reliability of the reference standard could have been improved if the study would have adopted an inter-judge approach. In the current study, only one expert opinion was used in isolation for every decision made regarding fitness to drive. Ideally, two experts would have independently evaluated the same patients and made their own decision regarding the patient's fitness to drive. Inter-judge agreement could have helped to increase the reliability of the reference standard. Unfortunately, due to the retrospective nature of the study, this was not possible. However, efforts were made to mitigate this effect by incorporating the clinical opinion of two experts. Despite having the inability to reach consensus on each individual decision, cut-off points on neuropsychological tests were established using two experts' clinical opinions. Subsequent studies should strive to validate the serial trichotomization method using different reference standards. As mentioned previously, each reference standard is comprised of its own advantages and disadvantages and validation against multiple reference standards has the potential of establishing more accurate cut-off points on common neuropsychological tests. Using a simulator as a reference standard could yield more precise cut-off points by including maneuvers that are

typically problematic in older adults with cognitive impairments. As stated previously by Gibbons and colleagues (2017), the serial tricotomization method should further be examined using a bigger sample size and a variety of different common neuropsychological tests in order to identify the best possible combination of tools to inform physicians in most decisions regarding fitness to drive among older adults with cognitive impairments.

Appendix

Appendix A.
Freund CDT Scoring Scale.

Time (3 points)	One hand points 2 (or symbol representative of 2) Exactly two hands *Absence* of intrusive marks, e.g., writing or hands indicating incorrect time, hand points to number 10; tic marks, time written in text (11:10; ten after eleven)
Numbers (2 points)	Numbers are inside the clock circle All numbers 1–12 are present, no duplicates or omissions
Spacing (2 points)	Numbers spaced equally or nearly equally from each other Numbers spaced equally or nearly equally from the edge of the circle

References

Adler, G., & Rottunda, S. J. (2011). The driver with dementia: A survey of physician attitudes, knowledge, and practice. *American Journal of Alzheimer's Disease & Other Dementiasr,26*(1), 58-64. doi:10.1177/1533317510390350

Abbott, K. V., Barton, F. B., Terhorst, L., & Shembel, A. (2016). Retrospective studies: A fresh look. *American Journal of Speech-Language Pathology,25*(2), 157-163. doi:10.1044/2016_ajslp-16-0025

Adler, G., & Kuskowski, M. (2003). Driving cessation in older men with dementia. *Alzheimer's Disease & Associated Disorders,17*(2), 68-71. doi:10.1097/00002093-200304000-00003

Aksan, N., Anderson, S. W., Dawson, J., Uc, E., & Rizzo, M. (2015). Cognitive functioning differentially predicts different dimensions of older drivers' on-road safety. *Accident Analysis & Prevention,75*, 236-244. doi:10.1016/j.aap.2014.12.007

Alzheimer Society of Canada. (2010). Rising tide: The impact of dementia on Canadian society. Retrieved from: https://alzheimer.ca/sites/default/files/files/national/advocacy/asc_rising_tide_full_report_e.pdf

Alzheimer Society of Canada (2017) A new way of looking at the impact of dementia in Canada. Retrieved from: https://alzheimer.ca/en/cornwall/Awareness/A-new-way-of-looking-at-dementia

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, S. W., Aksan, N., Dawson, J. D., Uc, E. Y., Johnson, A. M., & Rizzo, M. (2012). Neuropsychological assessment of driving safety risk in older adults with and without

neurologic disease. *Journal of Clinical and Experimental Neuropsychology,34*(9), 895-905. doi:10.1080/13803395.2011.630654

Anstey, K., Wood, J., Lord, S., & Walker, J. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review,25*(1), 45-65. doi:10.1016/j.cpr.2004.07.008

Apolinario, D., Magaldi, R. M., Busse, A. L., Lopes, L. D., Kasai, J. Y., & Satomi, E. (2009). Cognitive impairment and driving: A review of the literature. *Dementia & Neuropsychologia,3*(4), 283-290. doi:10.1590/s1980-57642009dn30400004

Aslaksen, P. M., Ørbo, M., Elvestad, R., Schäfer, C., & Anke, A. (2013). Prediction of on-road driving ability after traumatic brain injury and stroke. *European Journal of Neurology,20*(9), 1227-1233. doi:10.1111/ene.12172

Austroads (2004). Austroads research report: Model licence re-assessment procedure for older drivers: Stage 2 research (No.AP-R259/04). Sydney: Austroads.

Badenes, D., Garolera, M., Casas, L., Cejudo-Bolivar, J. C., Zaragoza, S., Calzado, N., & Aguilar, M. (2017). Relationship between neuropsychological tests and drivers license renewal tests in Parkinsons disease. *Traffic Injury Prevention,19*(2), 125-132. doi:10.1080/15389588.2017.1360491

Bang, J., Spina, S., & Miller, B. L. (2015). Frontotemporal dementia. *Lancet (London, England)*, *386*(10004), 1672-82. doi: 10.1016/S0140-6736(15)00461-4

Bédard, M., & Dickerson, A. E. (2014). Consensus statements for screening and assessment tools. *Occupational Therapy In Health Care,28*(2), 127-131. doi:10.3109/07380577.2014.903017

Bédard, M., Marshall, S., Man-Son-Hing, M., Weaver, B., Gélinas, I., Korner-Bitensky, N., . . . Vrkljan, B. (2013). It is premature to test older drivers with the SIMARD-MD. *Accident Analysis & Prevention,61*, 317-321. doi:10.1016/j.aap.2013.04.001

Belton-Chevallier, L., Motte-Baumvol, B., Fol, S., Jouffe, Y. (2017) Coping with the costs of car dependency : A system of expedients used by low-income households on the outskirts of Dijon and Paris. Retrieved from https://pdfs.semanticscholar.org/359e/f76bc7088c79d88a7ae097a62628a4eb35e4.pdf

Berndt, A., Clark, M., & May, E. (2008). Dementia severity and on-road assessment: Briefly revisited. *Australasian Journal on Aging,27*(3), 157-160. doi:10.1111/j.1741-6612.2008.00300.x

Berndt, A. H., May, E., & Darzins, P. (2015). On-road driving assessment and route design for drivers with dementia. *British Journal of Occupational Therapy,78*(2), 121-130. doi:10.1177/0308022614562397

Bott, N. T., Radke, A., Stephens, M. L., & Kramer, J. H. (2014). Frontotemporal dementia: Diagnosis, deficits and management. *Neurodegenerative Disease Management,4*(6), 439-454. doi:10.2217/nmt.14.34

Brandmaier, A. M., Ram, N., Wagner, G., & Gerstorf, D. (2017). Terminal Decline in Well-Being: The Role of Multi-Indicator Constellations of Physical Health and Psychosocial Correlates. doi:10.31234/osf.io/b9vd3

Brown, L. B., Ott, B. R., Papandonatos, G. D., Sui, Y., Ready, R. E., & Morris, J. C. (2005). Prediction of On-Road Driving Performance in Patients with Early Alzheimer's Disease. *Journal of the American Geriatrics Society,53*(1), 94-98. doi:10.1111/j.1532-5415.2005.53017.x

Burns, T., Lawler, K., Lawler, D., Mccarten, J. R., & Kuskowski, M. (2018). Predictive value of the cognitive performance test (CPT) for Staging Function and Fitness to Drive in People With Neurocognitive Disorders. *American Journal of Occupational Therapy,72*(4). doi:10.5014/ajot.2018.027052

Canadian Council of Motor Transport Administrators. (2017). Determining driver fitness in Canada. Retrieved from https://ccmta.ca/images/pdf-documents-english/CCMTA-Medical-Standards-2017-English.pdf

Carr, D. B., Barco, P. P., Wallendorf, M. J., Snellgrove, C. A., & Ott, B. R. (2011). Predicting road test performance in drivers with dementia. *Journal of the American Geriatrics Society,59*(11), 2112-2117. doi:10.1111/j.1532-5415.2011.03657.x

Chee, J. N., Rapoport, M. J., Molnar, F., Herrmann, N., Oneill, D., Marottoli, R., . . . Carr, D. B. (2017). Update on the risk of motor vehicle collision or driving impairment with dementia: A collaborative international systematic review and meta-analysis. *The American Journal of Geriatric Psychiatry,25*(12), 1376-1390. doi:10.1016/j.jagp.2017.05.007

Classen, S., Wang, Y., Crizzle, A. M., Winter, S. M., & Lanford, D. N. (2013). Predicting older driver on-road performance by means of the Useful Field of View and Trail Making Test Part B. *American Journal of Occupational Therapy,67*(5), 574-582. doi:10.5014/ajot.2013.008136

Contrada, R. J., Boulifard, D. A., Hekler, E. B., Idler, E. L., Spruill, T. M., Labouvie, E. W., & Krause, T. J. (2008). Psychosocial factors in heart surgery: Presurgical vulnerability and postsurgical recovery. *Health Psychology,27*(3), 309-319. doi:10.1037/0278-6133.27.3.309

Davis, R. L., & Ohman, J. M. (2016). Driving in early-stage Alzheimer's disease: An Integrative Review of the Literature. *Research in Gerontological Nursing,10*(02), 86-100. doi:10.3928/19404921-20160920-02

Davis, J., Wang, S., Festa, E., Luo, G., Moharrer, M., Bernier, J., & Ott, B. (2018). Detection of risky driving behaviors in the naturalistic environment in healthy older adults and mild Alzheimer's disease. *Geriatrics,3*(2), 13. doi:10.3390/geriatrics3020013

Dening, T., & Sandilyan, M. B. (2015). Dementia: Definitions and types. *Nursing Standard,29*(37), 37-42. doi:10.7748/ns.29.37.37.e9405

Diamond, A. (2013). Executive functions. *Annual Review of Psychology,64*(1), 135-168. doi:10.1146/annurev-psych-113011-143750

Dickerson, A. E., Molnar, L., Bedard, M., Eby, D. W., Classen, S., & Polgar, J. (2017). Transportation and aging: An updated research agenda for advancing safe mobility. *Journal of Applied Gerontology,* 073346481773915. doi:10.1177/0733464817739154

Dobbs, B. M., & Schopflocher, D. (2010). The introduction of a new screening tool for the identification of cognitively impaired medically at-risk drivers. *Journal of Primary Care & Community Health,1*(2), 119-127. doi:10.1177/2150131910369156

Dobbs, B. M., & Shergill, S. S. (2013). How effective is the Trail Making Test (Parts A and B) in identifying cognitively impaired drivers? *Age and Aging,42*(5), 577-581. doi:10.1093/aging/aft073

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*(10), 1342-1347. doi:10.1046/j.1532-5415.2003.51481.x

Ellaway, A., Macintyre, S., Hiscock, R., & Kearns, A. (2003). In the driving seat: Psychosocial benefits from private motor vehicle transport compared to public transport. *Transportation Research Part F: Traffic Psychology and Behaviour*,*6*(3), 217-231. doi:10.1016/s1369-8478(03)00027-5

Eshkoor, S. A., Mun, C. Y., Ng, C. K., & Hamid, T. A. (2015). Mild cognitive impairment and its management in older people. *Clinical Interventions in Aging*, 687. doi:10.2147/cia.s73922

Esser, P., Dent, S., Jones, C., Sheridan, B. J., Bradley, A., Wade, D. T., & Dawes, H. (2015). Utility of the MOCA as a cognitive predictor for fitness to drive. *Journal of Neurology, Neurosurgery & Psychiatry*,*87*(5), 567-568. doi:10.1136/jnnp-2015-310921

Evans, L. (2000). Risks older drivers face themselves and threats they pose to other road users. *International Journal of Epidemiology*,*29*(2), 315-322. doi:10.1093/ije/29.2.315

Fairhall, N., Sherrington, C., Cameron, I. D., Blyth, F., Naganathan, V., Clemson, L., . . . Cumming, R. G. (2014). Predicting participation restriction in community-dwelling older men: The Concord Health and Aging in Men Project. *Age and Aging*,*43*(1), 31-37. doi:10.1093/aging/aft111

Foley, D. J., Masaki, K. H., Ross, G. W., & Md, L. R. (2000). Driving cessation in older men with incident dementia. *Journal of the American Geriatrics Society*,*48*(8), 928-930. doi:10.1111/j.1532-5415.2000.tb06889.x

Fox, G. K., Bowden, S. C., Bashford, G. M., & Smith, D. S. (1997). Alzheimers disease and driving: Prediction and assessment of driving performance. *Journal of the American Geriatrics Society,45*(8), 949-953. doi:10.1111/j.1532-5415.1997.tb02965.x

Fraade-Blanar, L. A., Ebel, B. E., Larson, E. B., Sears, J. M., Thompson, H. J., Chan, K. C., & Crane, P. K. (2018). Cognitive decline and older driver crash risk. *Journal of the American Geriatrics Society,66*(6), 1075-1081. doi:10.1111/jgs.15378

Freeman, E. E., Gange, S. J., Muñoz, B., & West, S. K. (2006). Driving status and risk of entry into long-term care in older adults. *American Journal of Public Health,96*(7), 1254-1259. doi:10.2105/ajph.2005.069146

Freund, B., Gravenstein, S., Ferris, R., Burke, B. L., & Shaheen, E. (2005). Drawing clocks and driving cars. *Journal of General Internal Medicine,20*(3), 240-244. doi:10.1111/j.1525-1497.2005.40069.x

Frittelli, C., Borghetti, D., Iudice, G., Bonanni, E., Maestri, M., Tognoni, G., . . . Iudice, A. (2009). Effects of Alzheimers disease and mild cognitive impairment on driving ability: A controlled clinical study by simulated driving test. *International Journal of Geriatric Psychiatry,24*(3), 232-238. doi:10.1002/gps.2095

Gibbons, C., Smith, N., Middleton, R., Clack, J., Weaver, B., Dubois, S., & Bédard, M. (2017). Using serial trichotomization with common cognitive tests to screen for fitness to drive. *American Journal of Occupational Therapy,71*(2). doi:10.5014/ajot.2017.019695

Gold, D. A. (2012). An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology,34*(1), 11-34. doi:10.1080/13803395.2011.614598

Goldstein, G. & Watson, J.R. (1989). Test-retest reliability of the Halstead-Reitan Battery and the WAIS in a Neuropsychiatric Population. *The Clinical Neuropsychologist*, 3(3), 265-273.

Harada, C. N., Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*,*29*(4), 737-752. doi:10.1016/j.cger.2013.07.002

Hauer E. (1998) The safety of the older persons at intersections. Transportation in an aging society: Improving mobility and safety for Older persons- Volume 2; Special Report 218, pp. 194-252. Transportation Research Board, Washington, DC.

Herrmann, N. (2006). Predictors of driving cessation in mild-to-moderate dementia. *Canadian Medical Association Journal*,*175*(6), 591-595. doi:10.1503/cmaj.051707

Hird, M. A., Vetivelu, A., Saposnik, G., & Schweizer, T. A. (2014). Cognitive, on-road, and simulator-based driving assessment after Stroke. *Journal of Stroke and Cerebrovascular Diseases*,*23*(10), 2654-2670. doi:10.1016/j.jstrokecerebrovasdis.2014.06.010

Hoggarth, P. A., Innes, C. R., 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 & Prevention*,*42*(6), 1759-1768. doi:10.1016/j.aap.2010.04.017

Hollis, A. M., Duncanson, H., Kapust, L. R., Xi, P. M., & Oconnor, M. G. (2015). Validity of the Mini-Mental State Examination and the Montreal Cognitive Assessment in the Prediction of Driving Test Outcome. *Journal of the American Geriatrics Society*,*63*(5), 988-992. doi:10.1111/jgs.13384

Innes, C., Jones, R., Anderson, T., Dalrymple-Alford, J., Hayes, S., Hollobon, S., . . . Nicholls, A. (2005). Prediction of driving ability in persons with brain disorders using sensory-motor

and cognitive tests. *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. doi:10.1109/iembs.2005.1615713

Jang, R. W., Man-Son-Hing, M., Molnar, F. J., Hogan, D. B., Marshall, S. C., Auger, J., . . . Naglie, G. (2007). Family physicians' attitudes and practices regarding assessments of medical fitness to drive in older persons. *Journal of General Internal Medicine,22*(4), 531-543. doi:10.1007/s11606-006-0043-x

Jellinger, K., & Attems, J. (2007). Neuropathological evaluation of mixed dementia. *Journal of the Neurological Sciences,257*(1-2), 80-87. doi:10.1016/j.jns.2007.01.045

Joseph, P. G., Odonnell, M. J., Teo, K. K., Gao, P., Anderson, C., Probstfield, J. L., . . . Yusuf, S. (2014). The Mini-Mental State Examination, clinical factors, and motor vehicle crash risk. *Journal of the American Geriatrics Society,62*(8), 1419-1426. doi:10.1111/jgs.12936

Jutkowitz, E., Maclehose, R. F., Gaugler, J. E., Dowd, B., Kuntz, K. M., & Kane, R. L. (2016). Risk factors associated with cognitive, functional, and behavioral trajectories of newly diagnosed dementia patients. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences,72*(2), 251-258. doi:10.1093/gerona/glw079

Kaji, A. H., Schriger, D., & Green, S. (2014). Looking through the retrospectoscope: Reducing bias in emergency medicine chart review studies. *Annals of Emergency Medicine,64*(3), 292-298. doi:10.1016/j.annemergmed.2014.03.025

Kalaria, R. N. (2018). The pathology and pathophysiology of vascular dementia. *Neuropharmacology,134*, 226-239. doi:10.1016/j.neuropharm.2017.12.030

Kay, L. G., Bundy, A. C., & Clemson, L. M. (2009). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation,90*(9), 1514-1522. doi:10.1016/j.apmr.2009.03.011

Kay, L. G., Bundy, A. C., Clemson, L., Cheal, B., & Glendenning, T. (2012). Contribution of off-road tests to predicting on-road performance: A critical review of tests. *Australian Occupational Therapy Journal,59*(1), 89-97. doi:10.1111/j.1440-1630.2011.00989.x

Korczyn, A. D., Vakhapova, V., & Grinberg, L. (2012). Vascular dementia. *Journal of the Neurological Sciences,322*(1-2), 1. doi:10.1016/j.jns.2012.09.016

Kristoferitsch, W., Aboulenein-Djamshidian, F., Jecel, J., Rauschka, H., Rainer, M., Stanek, G., & Fischer, P. (2018). Secondary dementia due to Lyme neuroborreliosis. *Wiener Klinische Wochenschrift,130*(15-16), 468-478. doi:10.1007/s00508-018-1361-9

Kua, E. H., Ho, E., Tan, H. H., Tsoi, C., Thng, C., & Mahendran, R. (2014). The natural history of dementia. *Psychogeriatrics,14*(3), 196-201. doi:10.1111/psyg.12053

Larson, E. B., Yaffe, K., & Langa, K. M. (2013). New insights into the dementia epidemic. *New England Journal of Medicine,369*(24), 2275-2277. doi:10.1056/nejmp1311405

Ledger, S., Bennett, J. M., Chekaluk, E., & Batchelor, J. (2019). Cognitive function and driving: Important for young and old alike. *Transportation Research Part F: Traffic Psychology and Behaviour,60*, 262-273. doi:10.1016/j.trf.2018.10.024

Levenstein, S. (2002). Psychosocial factors in peptic ulcer and inflammatory bowel disease. *Journal of Consulting and Clinical Psychology,70*(3), 739-750. doi:10.1037/0022-006x.70.3.739

Lew, H. L., Poole, J. H., Lee, E. H., Jaffe, D. L., Huang, H., & Brodd, E. (2005). Predictive validity of driving-simulator assessments following traumatic brain injury: A preliminary study. *Brain Injury,19*(3), 177-188. doi:10.1080/02699050400017171

Macintyre, S., Hiscock, R., Kearns, A., Ellaway, A. (2001). Housing tenure and car access: Further exploration of the nature of their relations with health in a UK setting. *Journal of Epidemiology & Community Health,55*(5), 330-331. doi:10.1136/jech.55.5.330

Marshall, S., Demmings, E. M., Woolnough, A., Salim, D., & Man-Son-Hing, M. (2012). Determining Fitness to Drive in Older Persons: A Survey of Medical and Surgical Specialists. *Canadian Geriatrics Journal,15*(4). doi:10.5770/cgj.15.30

Marshall, S. C., Man-Son-Hing, M., Charlton, J., Molnar, L. J., Koppel, S., & Eby, D. W. (2013). The Candrive/Ozcandrive prospective older driver study: Methodology and early study findings. *Accident Analysis & Prevention,61*, 233-235. doi:10.1016/j.aap.2013.07.007

Manning, K., Davis, J., Papandonatos, G., & Ott, B. (2013). Clock drawing as a screen for impaired driving in aging and dementia: Is it worth the time? *Alzheimers & Dementia,9*(4). doi:10.1016/j.jalz.2013.05.966

Marshall, S., Molnar, F., Man-Son-Hing, M., Blair, R., Brosseau, L., Finestone, H., . . . Wilson, K. (2007). Predictors of driving ability following stroke: A systematic review. *Topics in Stroke Rehabilitation,14*(1), 98-114. doi:10.1310/v732-0916-m484-7314

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2015). Dropout during a driving simulator study: A survival analysis. *Journal of Safety Research,55*, 159-169. doi:10.1016/j.jsr.2015.08.004

Mathias, J. L., & Lucas, L. K. (2009). Cognitive predictors of unsafe driving in older drivers: A meta-analysis. *International Psychogeriatrics,21*(04), 637. doi:10.1017/s1041610209009119

Mayo, M., Bordelon, Y. (2014). Dementia with lewy bodies. *Seminars in Neurology, 34*(02), 182-188. doi:10.1055/s-0034-1381741

Maurer, K., Volk, S., Gerbaldo, H. (1997). Auguste D and Alzheimer's disease. *The Lancet*, (349), 1546-1549. Retrieved from https://alzheimer.neurology.ucla.edu/pubs/alzheimerLancet.pdf

Mayhew, D. R., Simpson, H. M., & Ferguson, S. A. (2006). Collisions involving senior drivers: high-risk conditions and locations. *Traffic Injury Prevention,7*(2), 117-124. doi:10.1080/15389580600636724

McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J. P., Weintraub, … Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, *89*(1), 88-100. doi: 10.1212/WNL.0000000000004058

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology,43*(11), 2412-2412. doi:10.1212/wnl.43.11.2412-a

Murman, D. (2015). The impact of age on cognition. *Seminars in Hearing,36*(03), 111-121. doi:10.1055/s-0035-1555115

National Association of Federal Retirees. (n.a). Provincial driving regulation for seniors. Retrieved on January 5th, 2018 from https://www.federalretirees.ca/en/My-Lifestyle/Driving

Nordlund, A., Rolstad, S., Klang, O., Lind, K., Hansen, S., & Wallin, A. (2007). Cognitive profiles of mild cognitive impairment with and without vascular disease. *Neuropsychology,21*(6), 706-712. doi:10.1037/0894-4105.21.6.706

O'Brien, J. T., & Thomas, A. (2015). Vascular dementia. *The Lancet, 386*(10004), 1698-1706. doi:10.1016/S0140-6736(15)00463-8

Olney, N., Spina, S., Miller, B. (2017). Frontotemporal dementia. *Neurologic Clinic*, 35(2), 339-374. doi.org/10.1016/j.ncl.2017.01.008

Ontario Ministry of Transportation. (2017). Senior driver's license renewal program. Retrieved from http://www.mto.gov.on.ca/english/driver/senior-driver-licence-renewal-program.shtml

Oswanski, M. F., Sharma, O. P., Raj, S. S., Vassar, L. A., Woods, K. L., Sargent, W. M., & Pitock, R. J. (2007). Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine & Rehabilitation,86*(3), 190-199. doi:10.1097/phm.0b013e31802b7de5

Ott, B. R., Heindel, W. C., Papandonatos, G. D., Festa, E. K., Davis, J. D., Daiello, L. A., & Morris, J. C. (2008). A longitudinal study of drivers with Alzheimer disease. *Neurology,70*(14), 1171-1178. doi:10.1212/01.wnl.0000294469.27156.30

Pannese, E. (2011). Morphological changes in nerve cells during normal aging. *Brain Structure and Function,216*(2), 85-89. doi:10.1007/s00429-011-0308-y

Patten, R. V., Britton, K., & Tremont, G. (2018). Comparing the Mini-Mental State Examination and the modified Mini-Mental State Examination in the detection of mild cognitive impairment in older adults. *International Psychogeriatrics*, 1-9. doi:10.1017/s1041610218001023

Pence, B. W., Miller, W. C., & Gaynes, B. N. (2009). Prevalence estimation and validation of new instruments in psychiatric research: An application of latent class analysis and sensitivity analysis. *Psychological Assessment,21*(2), 235-239. doi:10.1037/a0015686

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. *Journal of Internal Medicine,275*(3), 214-228. doi:10.1111/joim.12190

Petersen RC, Morris JC. Clinical features. In: Petersen RC, ed. *Mild Cognitive Impairment: Aging to Alzheimer's Disease*. New York: Oxford University Press, Inc., 2003; 15–40.

Qiu, C., Kivipelto, M., Strauss, E. (2009). Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. Dialogues in Clinical Neuroscience (11) 111-128. Retrieved from http://resolver.scholarsportal.info/resolve/12948322/v11i0002/111_eoadodasti

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*(1), 85-93. doi:10.5555/0894-4105.18.1.85.supp

Rocca, W. A., Mielke, M. M., Vemuri, P., & Miller, V. M. (2014). Sex and gender differences in the causes of dementia: A narrative review. *Maturitas,79*(2), 196-201. doi:10.1016/j.maturitas.2014.05.008

Sacuiu, S. (2016). Dementias. *Neuroepidemiology Handbook of Clinical Neurology,* 123-151. doi:10.1016/b978-0-12-802973-2.00008-2

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging,30*(4), 507-514. doi:10.1016/j.neurobiolaging.2008.09.023

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence,39*(4), 222-232. doi:10.1016/j.intell.2011.03.001

Sharma, S., Arain, Mathur, Rais, Nel, Sandhu, . . . Johal. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatment*, 449. doi:10.2147/ndt.s39776

Statistics Canada, Canadian Social Trends. (2007). Dependance on cars in urban neighbourhoods (11-008-XWE o.85 2008001). Retrieved from https://www150.statcan.gc.ca/n1/pub/11-008-x/2008001/article/10503-eng.htm

Statistics Canada, Rural and Small Town Canada Analysis Bulletin. (2008). Seniors in rural Canada (21-006-X). Retrieved from http://publications.gc.ca/collection_2008/statcan/21-006-X/21-006-x2007008-eng.pdf

Statistics Canada, The National Household Survey. (2011). Commuting to work (99-012-X2011003). Retrieved from https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-012-x/99-012-x2011003_1-eng.cfm

Statistics Canada, Canadian Social Trends. (2012). Profile of seniors' transportation habits (11-008-X). Retrieved from https://www150.statcan.gc.ca/n1/pub/11-008-x/2012001/article/11619-eng.htm

Statistics Canada, Canadian community Health Survey. (2016). Healthy aging findings (82-003-X). Retrieved from: https://www150.statcan.gc.ca/n1/pub/82-003-x/2010004/article/11391/findings-resultats-eng.htm

Statistics Canada, 2016 Census Population. (2016). A portrait of the population aged 85 and older in 2016 in Canada (98-304-X). Retrieved from https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016004/98-200-x2016004-eng.cfm

Statistics Canada, Public Health Agency of Canada. (2017). Dementia in Canada. Retrieved from https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/diseases-

conditions/dementia-highlights-canadian-chronic-disease-surveillance/dementia-highlights-canadian-chronic-disease-surveillance.pdf

Statistics Canada. (2018). Vehicle registration, 2018 (11-001-X). Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/180615/dq180615e-eng.htm

Shechtman, O., Awadzi, K. D., Classen, S., Lanford, D. N., & Joo, Y. (2010). Validity and critical driving errors of on-road assessment for older drivers. *American Journal of Occupational Therapy*,*64*(2), 242-251. doi:10.5014/ajot.64.2.242

Shively, S., Scher, A. I., Perl, D. P., & Diaz-Arrastia, R. (2012). Dementia resulting from traumatic brain injury. *Archives of Neurology*,*69*(10). doi:10.1001/archneurol.2011.3747

Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*,*15*(6), 548-561. doi:10.1002/1099-1166(200006)15:63.0.co;2-u

Smith, G. D., Shipley, M. J., & Rose, G. (1990). Magnitude and causes of socioeconomic differentials in mortality: *Health Inequalities*, 113-128. doi:10.2307/j.ctt1t8955q.17

Sonnen, J. A., Larson, E. B., Crane, P. K., Haneuse, S., Li, G., Schellenberg, G. D., . . . Montine, T. J. (2007). Pathological correlates of dementia in a longitudinal, population-based sample of aging. *Annals of Neurology*,*62*(4), 406-413. doi:10.1002/ana.21208

Stephens, B., Mccarthy, D., Marsiske, M., Shechtman, O., Classen, S., Justiss, M., & Mann, W. (2005). International older driver consensus conference on assessment, remediation and counseling for transportation alternatives: Summary and recommendations. *Physical & Occupational Therapy In Geriatrics*,*23*(2), 103-121. doi:10.1300/j148v23n02_07

Teng, E. (1987) The modified Mini-Mental State Test. Retrieved from: http://www.med.uottawa.ca/courses/CMED6203/Index_notes/3MS.pdf

Trick, L., & Caird, J. (2011). Methodological issues when conducting research on older drivers. *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*. doi:10.1201/b10836-27

Tuokko, H., Griffith, L. E., Simard, M., & Taler, V. (2016). Cognitive measures in the Canadian Longitudinal study on aging. *The Clinical Neuropsychologist,31*(1), 233-250. doi:10.1080/13854046.2016.1254279

Tuokko, H., & Hunter, F. (2002). Using "age" as a fitness-to-drive criterion for older adults. Ottawa, ON: Law Commission of Canada.

Unsworth, C. A., Baker, A., Lannin, N., Harries, P., Strahan, J., & Browne, M. (2018). Predicting fitness-to-drive following stroke using the occupational therapy – Driver Off Road Assessment Battery. *Disability and Rehabilitation,* 1-6. doi:10.1080/09638288.2018.1445784

Usher-Smith, J. A., Sharp, S. J., & Griffin, S. J. (2016). The spectrum effect in tests for risk prediction, screening, and diagnosis. *Bmj,* I3139. doi:10.1136/bmj.i3139

Wada-Isoe, K., Uemura, Y., Nakashita, S., Yamawaki, M., Tanaka, K., Yamamoto, M., . . . Nakashima, K. (2012). Prevalence of dementia and mild cognitive impairment in the rural island town of Ama-cho, Japan. *Dementia and Geriatric Cognitive Disorders Extra,2*(1), 190-199. doi:10.1159/000338244

Wadley, V. G., Okonkwo, O., Crowe, M., Vance, D. E., Elgin, J. M., Ball, K. K., & Owsley, C. (2009). Mild cognitive impairment and everyday function: An investigation of driving performance. *Journal of Geriatric Psychiatry and Neurology,22*(2), 87-94. doi:10.1177/0891988708328215

Wagner, J., Müri, R., Nef, T., & Mosimann, U. (2011). Cognition and driving in older persons. *Swiss Medical Weekly*. doi:10.4414/smw.2011.13136

Walshe, E., Mcintosh, C. W., Romer, D., & Winston, F. (2017). Executive function capacities, negative driving behavior and crashes in young drivers. *International Journal of Environmental Research and Public Health,14*(11), 1314. doi:10.3390/ijerph14111314

Wardlaw, J. M., Smith, C., & Dichgans, M. (2013). Mechanisms of sporadic cerebral small vessel disease: Insights from neuroimaging. *The Lancet Neurology,12*(5), 483-497. doi:10.1016/s1474-4422(13)70060-7

Weaver, B., Walter, S. D., & Bédard, M. (2013). How to report and interpret screening test properties: Guidelines for driving researchers. *Traffic Injury Prevention,15*(3), 252-261. doi:10.1080/15389588.2013.816692

Webster, N. (1963). New collegiate dictionary. A Merriam-Webster. Springfield, MA: G. & C. Merriam.

Wernham, M. L., Jarrett, P. G., Stewart, C., Macdonald, E., Macneil, D., & Hobbs, C. (2014). Comparison of the SIMARD MD to clinical impression in assessing fitness to drive in patients with cognitive impairment. *Canadian Geriatrics Journal,17*(2). doi:10.5770/cgj.17.100

Whishaw, I. Q., & Kolb, B. (2015). *Fundamentals of human neuropsychology*. New York, NY: Worth Custom Publishing.

Wilson, L. R., & Kirby, N. H. (2008). Individual differences in South Australian general practitioners' knowledge, procedures and opinions of the assessment of older drivers. *Australasian Journal on Aging,27*(3), 121-125. doi:10.1111/j.1741-6612.2008.00304.x

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

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*(6), 986-993. doi:10.1111/j.1532-5415.2008.01709.x

Zekry, D., Hauw, J., & Gold, G. (2002). Mixed dementia: epidemiology, diagnosis, and treatment. *Journal of the American Geriatrics Society,50*(8), 1431-1438. doi:10.1046/j.1532-5415.2002.50367.x

Zekry, D., Duyckaerts, C., & Hauw, J. (2007). Démences vasculaires et démences mixtes. *La Presse Médicale,36*(10), 1469-1476. doi:10.1016/j.lpm.2007.04.020

Zhang, J., Fraser, S., Lindsay, J., Clarke, K., Mao, Y. (1998). Age-specific patterns of factors related to fatal motor vehicle traffic crashes focus on young and elderly drivers. *Public Health,112*(5), 289-295. doi:10.1016/s0033-3506(98)00257-1