SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

3392.

SZABÓ GÁBOR

A támasztó és mozgató szervrendszer működésének fiziológiája

című program

Programvezető: Dr. Szőke György, egyetemi tanár

Témavezető: Dr. Fazekas Gábor, egyetemi tanár

COGNITIVE PREDICTORS OF FITNESS TO DRIVE: CLINICAL EVALUATION IN STROKE AND OLDER ADULTS

PhD thesis

Gábor Szabó

Semmelweis University Doctoral College

Surgical Medicine Division

Supervisor:	Gábor Fazekas, MD, Ph.D
Official reviewers:	Cecília Varjú, MD, Ph.D
	Ildikó Sipos, MD, Ph.D

Head of the Complex Examination Committee:	György Szőke, MD, D.Sc
Members of the Complex Examination Committee:	Áron Lazáry, MD, Ph.D
	Gábor Skaliczki, MD, Ph.D
	Imre Szerb MD, Ph.D

Budapest

2026

Table of contents

List of abbreviations

ACC	Accuracy
AUC	Area Under the Curve
DHMT	Drive Home Maze Test
IES	Inverse Efficiency Score
LOOCV	Leave-One-Out Cross-Validation
MCI	Mild Cognitive Impairment
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
OT-DORA	Occupational Therapy Driver Off-road Assessment Battery
ROC	Receiver Operating Characteristic
ROC-AUC	Receiver Operating Characteristic – Area Under the Curve
RLRCT-H	Road Law and Road Craft Test – Hungarian version
RT	Reaction Time
SCT	Sensorimotor Coordination Test
SCWT	Stroop Color and Word Test
SDSA	Stroke Drivers Screening Assessment
SHAP	SHapley Additive exPlanations
SNT	Starry Night Test
TMT-A	Trail Making Test, Part A
TMT-B	Trail Making Test, Part B
WHO	World Health Organization

1. Introduction

Driving has become an essential part of modern life. For many individuals, it provides independence and serves as an asset for social integration, access to healthcare, shopping, and commuting to work.[1] At the same time, it has become a symbol of freedom, success, and personal competence. For these reasons, reliance on motor vehicles remains strong, even among risk groups such as older adults and stroke patients with significant cognitive symptoms that might interfere with safe driving.[2,3]

Every year, more than 1.3 million individuals die and 50 million are severly injured worldwide as a result of road traffic crashes.[4] Consequently, traffic accidents are the leading cause of death among those aged 5–29 years. More than half of the victims are vulnerable road users, such as pedestrians or cyclists. Hungarian statistics reflect a similar pattern. Police reports from 2024 document 14,687 road collisions, including 446 fatal cases and 4,143 serious injuries.[5] In response to such figures, the World Health Organization introduced the Global Plan for the Decade of Action for Road Safety (2021–2030), aiming to reduce road traffic deaths by half.[6] This initiative later received political endorsement when the United Nations General Assembly proclaimed the "Second Decade of Action for Road Safety". As part of this complex intervention plan, the WHO strives for strengthening safe road user behaviour in order to reduce future fatalities and injuries. As more than two thirds of car accidents are caused by human error, it is crucial to understand the fundamental factors influencing driver behaviour.[7]

1.1. Models of driving behaviour

In the previous decades, several theoretical models made attempts to explain driving behaviour, but none has offered a complete solution. Early work focused on relatively stable individual traits, such as the idea of "accident proneness," the rather vague notion that some people are more likely to cause crashes regardless of circumstances.[8] To determine accident proneness, researchers examined vision, reaction times, and even personality factors, yet the correlations with crash involvement remained weak. These approaches were also limited by the fact that accidents are relatively rare and strongly influenced by situational factors. Later, motivational models suggested that drivers regulate their behaviour according to perceived risk. For example, Wilde's risk

homeostasis theory states that drivers try to keep their level of perceived risk at a constant level.[9] In contrast, Näätänen and Summala's zero-risk model suggested that most of the time drivers have a false sense of safety while driving, until a threshold is suddenly crossed.[10] These ideas broadened the scope of driving research, but they proved difficult to confirm in real-world settings, and the line between individual traits and environmental conditions often became hard to draw.

A more enduring contribution arose from cognitive psychology, through its integration with hierarchical control theory. Michon's hierarchical control model, introduced in 1985, remains one of the most influential and most cited models to this day.[11] It describes driving as a complex and partly overlearned everyday task organised hierarchically across three levels. At the strategic level, the driver makes broad decisions such as planning the trip, selecting specific routes, and avoiding night-time driving and known hazardous routes. The tactical level involves short-term manoeuvres, including overtaking, crossing hazardous intersections, or making subtle adjustments while driving in poor weather or low visibility. Finally, the operational level is usually automatic and refers to the immediate control of the vehicle, such as steering, braking, or gear shifting. According to Michon, these levels operate simultaneously and flexibly: when unexpected events occur, control may shift upward from automatic skills to more deliberate, problem-solving processes, making actions which are normally unconscious at the operational level become conscious. This makes the model particularly useful for understanding vulnerable groups: while ageing drivers often compensate by adapting their behaviour at all three levels, stroke patients may experience more specific deficits ranging from impaired spatial attention during manoeuvres to reduced executive control when making route choices.

Michon's framework integrates earlier theories by combining long-term goals, situational choices, and momentary control into a single structure. It also aligns well with core ideas from cognitive psychology, including attention and automatic, overlearned processes. Because of this, the model has become a practical basis not only for research but also for applied assessments, and it remains relevant today.[10]

1.2. Driving and ageing

Driving abilities of the elderly is an increasingly significant issue in modern societies.[12] With the growing number of older drivers on the road, many express concerns regarding their safety. As György Nemes points out, traffic accidents involving older drivers often receive excessive media attention, and the general public often demands strict regulations on older drivers.[13] However, such decisions have a significant impact on the affected person's life. According to a recent meta-analysis of Chihuri, losing driver's license is an independent risk factor for social isolation and cognitive decline, while the risk of depression is almost two-fold in this group.[14] Conclusively, decisions regarding driving fitness should always be based on evidence-based assessments.

1.2.1. Driving skills in physiological ageing and in neurocognitive disorders

In light of the worsening traffic safety statistics among drivers over the age of 70, it is important to emphasize that older adults do not constitute a homogeneous group.

1.2.1.1. Physiological ageing and driving

A certain degree of cognitive decline is a natural consequence of physiological ageing: selective and divided attention, visuospatial skills, and executive functions tend to deteriorate over time.[15,16]

As a result, both the quality and the speed of information processing may be affected. Resistance to distracting stimuli decreases, and due to psychomotor slowing, performance under time pressure may lead to slower and less optimal decisions.

The underlying cause of these changes is typically age-related gradual cortical atrophy.[17] Alongside physiological ageing, the volume of white matter decreases over time. This process takes place earlier and to a greater extent compared to grey matter loss.[18] These white matter changes are not evenly distributed across brain regions, with the most pronounced decline observed in the prefrontal areas.[19] With advancing age, signal transmission between brain regions may slow down in milder cases, or cease altogether in more severe instances - a phenomenon known as cortical disconnection. As a consequence, cognitive functions that rely on communication between distributed brain regions may become functionally impaired, and in some cases, psychiatric symptoms may emerge.

The progression of the neurobiological processes described above shows considerable individual variability. These changes are not necessarily pathological. Therefore, it is recommended to distinguish the driving performance of older adults who follow the physiological course of ageing from those with mild cognitive impairment or dementia.

According to a summary by Sieren and Haustein, chronological age alone is a poor predictor of driving fitness.[20] One explanation for this is that older individuals are generally aware of their changing cognitive abilities and tend to compensate for them. Compensation strategies may include avoiding situations that exceed their capabilities or maintaining a slower driving pace. The latter approach is successfully adopted by a significant proportion of older drivers, as it effectively increases the time available for information processing and the execution of appropriate actions. However, it is important to note that maintaining a low driving speed may make other road users do risky manoeuvres, thereby indirectly contributing to accidents.

In everyday situations, drivers are often forced to engage in complex situations where compensatory strategies cannot be applied. Dobbs, Heller, and Schopflocher specifically assessed driving performance in scenarios that place a high demand on older adults' reduced information-processing capacity, such as navigating intersections, changing lanes, or making unprotected left turns.[21] In their study, 3% of younger drivers, 25% of healthy individuals over the age of 65, and approximately 68% of those showing early clinical signs of cognitive decline were found to be unfit to manage such complex situations. Among the serious, safety-critical errors observed during test drives, 50% occurred during lane changes and another 21% during left turns. These findings were supported by a study by Lombardi, Horrey, and Courtney, who analyzed data from 120,000 fatal car crashes.[22] They found that while drivers under the age of 65 were responsible for 38% of fatal intersection-related accidents, this proportion was significantly higher (56%) among drivers aged 65 and older.

1.2.1.2. Age related neurocognitive disorders and driving

The practice of age-dependent screening in the entire population is not strongly supported by scientific evidence. However, there is broader consensus regarding the driving abilities of individuals with mild or major neurocognitive disorders.[23]

In the case of Mild Cognitive Impairment (MCI), standardized neuropsychological tests show significant cognitive deficits, but the individual can still live an independent life, as symptoms are not interfering with everyday activities.[24] The symptoms may progress to dementia at a rate of 5–10% per year, however, it is important to note that in most cases, the symptoms do not worsen even over a 10-year period.[25]

Current recommendations suggest that people with MCI should generally be considered fit to drive.[23,26] However, regular neuropsychological monitoring is highly recommended, and if symptoms worsen, the affected individual should be prepared to give up driving. If symptoms exceed the level of MCI and a major neurocognitive disorder is diagnosed, permanent license revocation may become necessary. Even mild dementia, based on the Clinical Dementia Rating scale, is linked to significantly reduced driving ability and increased risk of traffic accidents.[27] People with dementia are about ten times more likely to fail a driving test compared to healthy older adults.

More than 60% of individuals with neurocognitive disorders still drive, but the proportion decreases as cognitive impairment progresses.[28] Among those with very mild dementia (CDR-0.5), 46% are still active drivers, while in the case of mild dementia (CDR-1), it drops to 22%.

A major difficulty is that about 80% of people with dementia lack sufficient insight into their condition and therefore often disagree with being restricted from driving.[29] In more severe cases, they may not even remember being informed about the restriction. In one study, 172 out of 198 participants with dementia were still actively driving.[3] According to self-reports, 51% considered their driving abilities adequate, while only 29% of their relatives agreed with that assessment.

1.3. Driving after stroke

Every year almost 12 million new stroke cases are registered, and one in four people above the age of 25 years is expected to suffer a stroke.[30] With the survival rate improving, a growing number of patients are expressing their wish to return to driving. However, only one out of ten patients receives formal assessment after stroke.[2] According to a study conducted in Saudi Arabia, out of 94 stroke survivors, none received driving evaluation.[31]

Physical impairments such as hemiparesis are common after stroke. Since they are more visible and evidently interfere with everyday activities, their assessment is generally easier.[32] In many cases, physical impairments can also be managed with vehicle adaptations.[33] On the other hand, almost 40% of stroke patients present significant cognitive deficits, such as problems in attention, executive function, memory, metacognition, and visuoperceptual skills.[34] Even subtle deficits in any of these domains can render a patient unfit to drive, and their assessment is challenging for clinicians due to the lack of evidence-based cut-off scores.

1.3.1. Post-stroke cognitive deficits and driving abilities

Mild cognitive changes after stroke are often difficult to judge in standard neuropsychological tasks.[35] The same impairment becomes more noticeable when decisions must be made quickly, without the pauses and step-by-step processing that clinical tests usually allow. The stroke literature comparing patients with age-matched healthy drivers consistently shows that the key difference is not in accuracy or knowledge, but in speed of integration and allocation of attention under simultaneous demands. Patients often perform the same tasks as healthy drivers do, only too slowly for those actions to remain safe once translated to a moving vehicle.[36]

Attentional disturbances illustrate this mismatch. In research comparing stroke survivors with healthy peers, performance on visual attention tasks was not necessarily incorrect, but delayed, and relevant cues were more likely to lose priority when the visual field became crowded.[35] In real driving, gradual changes in speed or peripheral movement require sustained engagement; if the driver perceives the change only after a delay, the decision that follows may already be out of sync with the situation. These lapses differ from visual neglect or sensory loss, as they do not concern absence of perception, but loss of temporal adequacy.

Divided attention exposes an additional difficulty.[37] Laboratory comparisons show that stroke survivors can handle individual subtasks, but the moment two or more cognitive operations compete - such as planning a turn while monitoring oncoming traffic - the coordination becomes sequential rather than parallel. Driving, however, offers no sequential mode: perception, selection, and action must occur together. This temporal stacking error explains why drivers after stroke may exhibit technically correct

manoeuvres at inappropriate moments, producing risk despite possessing adequate procedural knowledge.

Executive abilities influence how drivers balance speed and caution.[38,39] Studies using both on-road evaluation and cognitive measures demonstrate that poor inhibitory control or inflexible planning may lead to disproportionate hesitation or premature decisions, neither of which stems from misunderstanding the rules of driving. A hesitant driver may obstruct the flow of traffic and provoke unpredictable responses from others; a hasty driver may fail to inhibit an earlier intention once conditions change. The behavioural variability itself becomes a hazard, since driving requires not merely correct decisions, but timely and stable decision-making patterns.[40]

Visuospatial deficits further complicate safety, not necessarily through object recognition problems, but through impaired estimation of spatial relations.[41] Stroke cohorts perform comparably to healthy drivers on object identification, yet they struggle to interpret where those objects are in relation to his own vehicle. On-road studies show that drivers compensate through reduced speed and increased caution. The necessity of these compensatory behaviours indicates that the constantly changing environment is harder for the patient to interpret.

Memory and metacognition influence not only keeping track of instructions, but how driving plans remain active and inform new actions.[42,43] Short-term memory deficits can leave recent events "unavailable" when needed, while impaired procedural memory forces drivers to consciously manage tasks that normally run automatically. Conscious control of simple manoeuvres consumes attentional resources normally used for hazard perception, and poor insight prevents drivers from compensating by altering habits, accepting restrictions, or seeking evaluation.

Stroke survivors do not necessarily fail because of a single cognitive deficit. The difficulty more often lies in how separate operations work together when decisions need to be made quickly.[35,36] A person may perform adequately in simple or predictable situations but struggle as soon as several tasks have to be coordinated at once. This is also why fixed cut-off points from isolated cognitive tests often tell us little about driving ability: they describe performance under controlled conditions, not how cognition behaves when it has to act rapidly and continuously.

1.3.2. Hemispatial neglect syndrome and driving abilities

Another striking example of post-stroke cognitive impairments is hemispatial neglect, a common consequence of right hemisphere stroke. It is characterised by a failure to attend to or respond to stimuli on the contralesional side of space, most often the left.[44,45] Unlike visual field loss, neglect is not a sensory problem but a disorder of spatial attention and representation: patients are able to see, yet part of their environment is not processed or acted upon. The syndrome is heterogeneous, and its clinical expression varies substantially across patients.

Within the visual domain, two patterns are typically distinguished.[46] Egocentric neglect involves a spatial bias defined with respect to the patient's own midline; information on one side of the environment is systematically overlooked. In a driving context, this may lead to failing to notice traffic, pedestrians or road markings on the neglected side. Allocentric neglect, by contrast, affects the perception of individual objects: one side of an object is consistently omitted or misinterpreted. For drivers, this can interfere with reading traffic signs, interpreting dashboard information or recognising hazards in time. Other forms are also documented, including personal neglect,[47] which affects perception of one side of the body, and motor neglect,[48] characterised by reduced intentional movement toward the contralesional side despite preserved strength. These variations are clinically important, as they influence steering control, lane maintenance and the effective use of mirrors.

Neglect is frequently multimodal, extending beyond vision to proprioceptive and tactile domains.[49] As a result, some patients misjudge limb position or vehicle trajectory, even if standard cancellation tests appear normal. Reduced insight is also common, meaning that affected individuals may underestimate their difficulties. Together, these features make hemispatial neglect one of the most important risk factors in the assessment of driving after stroke. Although there are nearly 300 different assessment tools around, most of them fall short in their validity and reliability.[44,50]

1.3.2.1. Computer-based assessment of hemispatial neglect

Given how strongly neglect can interfere with safe driving, there is a clear need for more sensitive ways of detecting it. In recent years, computer-based, dynamic methods have increasingly been recognised as a promising solution.[51] Initial evidence already suggested that these tasks are capable of revealing residual neglect that remains undetected by conventional paper-and-pencil tools, especially in patients who have reached the subacute or chronic stage and appear "recovered" on standard tests. They achieve this by minimising compensatory strategies, providing unpredictable stimulus presentation and capturing subtle delays in lateralised reaction times. Over the past decade, this approach has developed considerably.[52] Large-scale reviews demonstrate that modern computer-based tests outperform traditional assessment in terms of sensitivity, particularly for mild or chronic neglect presentations. Instead of reducing performance to a single score, these tools follow how the person responds throughout the task, measuring precision, reaction time, and fluctuations in attention. Looking at performance in this way offers a clearer picture of spatial attention and can highlight specific patterns of neglect, which may have obvious implications for everyday behaviour, including driving.[53] Moreover, systematic evidence indicates that these tests limit learning and compensation effects that frequently normalise paper-and-pencil scores despite persistent real-world impairment. By deploying brief, lateralised stimuli under high attentional load, computer-based paradigms can expose omissions that only emerge under complex task demands, similarly to real traffic environments. This methodological feature parallels the multi-tasking and rapid monitoring drivers require when scanning mirrors, reacting to hazards, or maintaining lane position in dense traffic.[36] While no universally accepted technological "gold standard" yet exists, recent reviews highlight that dynamic computer-based assessment provides richer data, higher sensitivity and better ecological validity than the tests most commonly applied in clinical stroke settings.[51] These advances underline the importance of integrating such methods into driving fitness evaluations, particularly when neglect is mild, fluctuating, or chronic.

1.3.3. Limitations of applying ageing research to stroke patients

Research evidence from the broader field of older drivers cannot be directly applied to stroke survivors. The symptoms of stroke patients vary significantly depending on the extent and site of the injury, making them a heterogeneous group requiring individual assessment.[54] The time course of stroke patients also differs from that of older drivers. Driving abilities of the elderly, especially in the presence of neurocognitive decline, deteriorate with age, indicating the need for regular medical and neuropsychological screening.[22] Time plays an inverse role in the case of stroke patients, as there is a linear correlation between time elapsed since the stroke and a positive outcome on the on-road test.[55] These results indicate that the timing of evaluation is critical in stroke patients, and premature assessment should be avoided.

1.4. Assessment of fitness to drive

Driving is a hazardous activity since even a healthy person carries a certain level of risk of causing an accident. Therefore, the purpose of driving assessments for stroke patients or the elderly is not to eliminate risk entirely, but rather to ensure that their risk level remains comparable to that of the general population.[56] Since psychometric testing is relatively cheap and accessible, research has long focused on finding and developing the ideal test to predict fitness to drive.[57] However, the field faces serious methodological problems.

1.4.1. Lack of outcome measures

The first issue is the lack of an objective outcome variable.[58] Although long term participation in real traffic would offer the most reliable evidence of fitness to drive, this practice is not viable for safety reasons. Even though driving simulators could provide ecologically valid information without safety concerns, they are not considered a first-line assessment.[35] The reasons include high cost, the common occurrence of simulator sickness, and the observation that participants often take more risks in a simulator than they would in real-world driving.[59,60]

At present, on-road tests are considered the gold standard assessment. Long-term follow-up has shown that patients who pass an on-road test have the same level of accident involvement as the general population. Despite their widespread use, on-road assessments

have many limitations as they are costly, highly situational and dependent on the actual traffic environment.[61] Furthermore, the presence and involvement of the evaluator may influence driving behaviour. In driving research, psychometric testing is often applied to predict the outcome of the on-road test.[62] Despite these efforts, the accuracy of these tests or test batteries remain relatively low, and false classifications remain a common problem.[58]

1.4.2. The issue of high misclassification rate

Canadian geriatrician Frank Molnar identified two main methodological problems in his 2006 systematic review.[58] According to Molnar, the first common issue in driving research is that studies often analyse the mere correlation between cognitive test results and the on-road test, but do not provide clinically useful cut-off points and other relevant psychometric properties, such as specificity and sensitivity. The second problem is the fact that test results of those who are fit to drive significantly overlap with the group that is unfit to drive.[63] As a consequence, determining one specific cut-off point would always result in a certain level of misclassification. If the clinician takes a cautious approach, the cut-off point can be adjusted accordingly. In this case, the model's sensitivity increases, but its specificity decreases. As a result, the model identifies unfit drivers with high confidence, but at the same time misclassifies many safe drivers as unfit (see Figure 1.).[64]
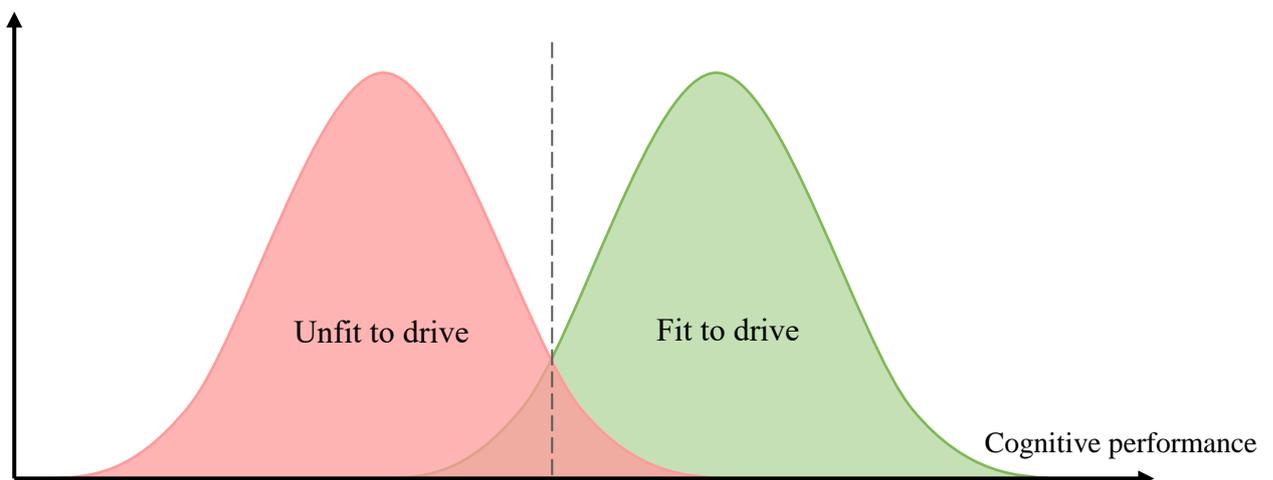
Figure 1. Classification using dichotomization based on a single cut-off value.[64]

To resolve this issue, trichotomization has been proposed. Instead of relying on a single cut-off point (i.e., dichotomization), two cut-off points are introduced, creating a third group usually referred to as inconclusive (see Figure 2.).[65] This method allows the clinician to avoid making an unfounded decision. Patients in the inconclusive group are generally a devised to undergo further evaluation, preferably an on-road test.
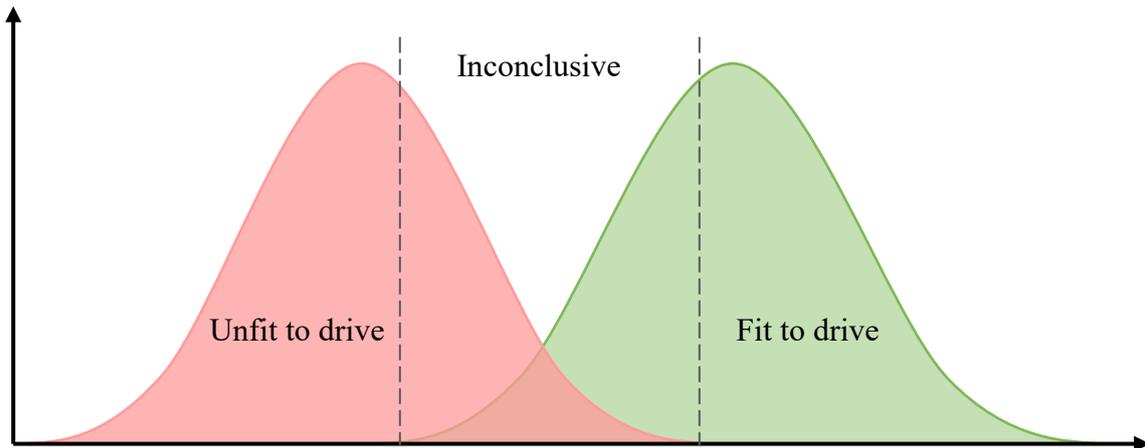
Figure 2. Trichotomous classification using two cut-off values; cases within the indeterminate zone require additional evaluation.[64]

1.4.3. Overfitting

Since Molnar's proposal, relevant advances have been made in the field of driving research, as studies have increasingly reported cut-off points, and multiple attempts have been made to use trichotomization. Results show that the proportion of patients in the inconclusive group varies between 17–36%.[63,65–67] However, with the emerging field of machine learning, a new issue was identified: the risk of overfitting.[68] Overfitting occurs when a model aligns perfectly with its training data, but fails to generalize to a new sample. The case of the Stroke Driver Screening Assessment (SDSA) provides an illustrative example.[69] The SDSA is composed of four cognitive tests and is widely used to assess post-stroke driving abilities. In the original sample of 40 British patients, the model classified correctly in over 80% of the cases. In later adaptations, the model was tested on novel samples. In its Swedish adaptation, the correct classification rate was only 62% in stroke cases and 50% in the elderly group with dementia or cognitive deficits.[70]

To avoid the pitfall of overfitting, rigorous validation techniques should be used, such as retesting in independent samples. Proper validation can also be achieved by applying resampling methods, including bootstrapping or leave-one-out cross-validation.[71] Other strategies to reduce overfitting include removing intercorrelated variables, increasing the ratio of participants to variables, or avoiding automatic variable selection. According to Hoggarth et al., many studies published in driving research either fail to use any technique to reduce overfitting, or do not provide enough detail to assess the risk of it.[72] Since the use of validation techniques can significantly reduce a model's accuracy and other psychometric properties, reported accuracy rates can be inflated. Recent recommendations emphasize that only classification rates after validation should be reported, and sufficient information about the method of variable selection should also be disclosed.

2. Objectives

Determining fitness to drive is a practical issue affecting many neurological patients worldwide. In modern societies, driving represents independence and access to daily activities, but for many patients it may also pose a considerable danger to themselves and other road users. Physical impairments are often more evident and easier to assess, while cognitive deficits are less apparent and generally more difficult to assess. For this reason, there is a need for methods that help clinicians make safer and more well-founded decisions. This dissertation was designed with that aim in mind.

The specific objectives were as follows:

1. **Introducing a new method of data analysis**
   Our goal was to build a logistic regression model validated with leave-one-out cross-validation. A key innovation was the use of trichotomization, by adding a third, uncertain category. We assumed that this approach could lower the rate of misclassification and support safer decision-making in clinical practice. While previous work has demonstrated the conceptual value of trichotomization, our study aims to extend this approach by evaluating a fully validated model, allowing its performance to be interpreted with greater methodological confidence.

2. **Investigating the effect of hemispatial neglect on driving performance**
   Neglect is a common post-stroke syndrome that can have serious consequences in traffic. Our aim was to examine how neglect influences driving skills and to what extent it is reflected in the outcome of the on-road test, as neglect is expected to negatively affect critical aspects of driving such as hazard detection, lane positioning, and spatial judgement.

3. **Evaluating the predictive value of traditional neglect tests**
   We analyzed whether clinical diagnoses of neglect (based on the Line Bisection Test or the Bells Test) are consistently associated with poorer performance during test drives. This may help clarify how much the field can rely on these classical tools in practice.

4. **To assess the utility of the Starry Night Test**
   The Starry Night Test is a dynamic, computer-based task that can detect subtle signs of neglect. The objective was to determine whether this test adds meaningful

value to the predictive model and can support decisions on fitness to drive, as the SNT specifically minimises compensatory strategies and captures lateralised reaction-time delays that may remain undetected in conventional paper-and-pencil tests.

5. **Providing recommendations for the Hungarian context**

Since there is currently no standardized protocol for cognitive aspects of driving assessment in Hungary, another aim was to summarize relevant international findings and translate them into practical guidance for clinicians and decision-makers.

6. **To explore possibilities for wider implementation**

Finally, the dissertation also considered whether the method tested in stroke patients could be adapted for use in other populations. In particular, a short trichotomized screening tool at the level of general practitioners could serve as a first-line filter before referral to specialist evaluation or on-road testing.

3. Methods

3.1. Narrative review

In order to inform Hungarian clinicians of the current state of the literature, a narrative review was conducted with the aim of summarizing and critically discussing the literature in a broad manner.[64] Given the heterogeneity of the studies and the lack of standardized outcome measures, a systematic review was deemed less feasible and less practical for this purpose. The topic of our review was the driving abilities of the elderly and the possibilities of their cognitive evaluation.

The literature search was carried out between November 2023 and March 2024 using PubMed, Scopus, and Web of Science, complemented by searches in Google Scholar for additional literature. In addition, Hungarian publications were identified through MATARKA database and using Google Scholar.

The search was performed in both English and Hungarian, using combinations of the following keywords: „Elderly", „Aged", „Automobile Driving", „Fitness to drive", „Driving assessment", „Cognitive deficits", „Neuropsychology" (English), and „Időskor", „Gépjárművezetés", „Vezetésre való alkalmasság", „Kognitív zavarok", „Neuropszichológia" (Hungarian). Both review articles and original studies were considered eligible. Publications were included if they addressed driving abilities of elderly drivers, the cognitive and neuropsychological aspects of elderly drivers, or clinical and methodological issues related to fitness-to-drive assessments. No restrictions were applied regarding study design or publication year, although only peer-reviewed sources were included.

Across all databases, 53 papers were found to be relevant to the topic. Most originated from international sources, and only two were published in Hungarian. Since the review focused specifically on studies examining driving in older adults or on cognitive approaches to driving assessment, no articles meeting these criteria required exclusion.

During the review, two lines of inquiry became especially important. A recurrent issue was whether chronological age, by itself, shows a demonstrable relation to driving ability, as opposed to being treated as a proxy for reduced safety. The literature increasingly

suggests that age alone is a weak predictor and that decisions should rely on a person's functional cognitive profile rather than their years.

The second issue concerned international practice: which cognitive tools are used abroad to evaluate driving, how they are interpreted, and how well their results reflect real driving behaviour. Placing these results in a common framework served to map international practice and highlight elements that could later inform evidence-based decision-making in Hungary, where validated assessment procedures are still limited.

3.1.1. Methodological implications in the Hungarian context

Research on fitness-to-drive assessment in older adults has expanded considerably over the past decades, particularly in international contexts.[73] When the focus is restricted to the Hungarian scientific literature, however, our narrative review reveals a substantial gap, as no empirical study was identified that examined the cognitive fitness to drive of older adults, regardless of the applied methods or outcome measures.[64]

Hungarian publications addressing fitness to drive among the elderly typically approach the topic from related but indirect perspectives, including traffic statistics, legal context, or general medical fitness for driving. Studies that focus specifically on the assessment of cognitive fitness to drive, or that connect cognitive test performance to real-world driving behaviour, are largely absent. As a consequence, cognitive screening practices currently applied in clinical and primary care settings operate without an empirically grounded methodology that would allow cognitive test results to be meaningfully linked to functional driving outcomes.[74]

The absence of empirical evidence has important methodological consequences.[72] Without validation against functional criteria, the interpretation of cognitive test results remains uncertain, particularly in cases that fall close to conventional decision thresholds.[64] The narrative review therefore suggests that the primary limitation of the current Hungarian approach lies not in the availability of cognitive assessments themselves, but in the lack of empirically supported decision criteria and structured decision-making frameworks.

3.2. Original study

This prospective observational study was carried out at the Rehabilitation Clinic of Semmelweis University, Budapest.[75] Participant recruitment started in January 2022 after receiving ethical approval from the Hungarian Medical Research Council (IV/649-1/2022/EKU) and continued until November 2024. The study is registered at ClinicalTrials.gov under the identifier NCT06111989. Participants were recruited from the Post-Stroke Rehabilitation Department and the Traumatic Brain Injury Department of the Rehabilitation Clinic.

3.2.1. Participants

Over the recruitment period, a total of 830 inpatients were screened according to the predefined inclusion and exclusion criteria. Patients with either ischemic or hemorrhagic stroke were enrolled, provided that at least six months had elapsed since the event and the patient held a valid driver's licence. Exclusion criteria included ophthalmological conditions, epilepsy, comprehension-affecting aphasia, and severe psychiatric disorders. Patients were also not eligible if they were physically unable to perform basic driving-related tasks (such as pressing pedals, steering, or maintaining sitting balance), or if an independent rehabilitation physician explicitly contraindicated driving because of physical disability.

From the 830 individuals screened, participation in the study was offered to all inpatients meeting the criteria. Altogether, 149 patients were invited, and 121 consented to take part, giving a participation rate of just over 80%. After accounting for six dropouts (three individuals were found not to hold valid driving licences, two declined to proceed with the on-road test after completing cognitive assessment, and one experienced an epileptic seizure before the on-road test), the final sample comprised 115 participants (mean age 52.9 years, SD 7.8; 77% male). Among them, 86 had suffered an ischemic stroke and 29 a hemorrhagic stroke. Written informed consent was obtained from every participant.

3.2.2. Procedure

Patient data were first collected from the medical records, including stroke type and location, time since onset, as well as age and sex. At study entry, patients also reported their educational background and previous driving experience.

Neglect was diagnosed primarily through bedside neurological examination and confirmed with neuropsychological tests. The latter included the Line Bisection Test, where an ipsilesional deviation of more than 9.5% from the center was taken as indicative of neglect,[76] and the Bells Test, where an Asymmetry Score of ≥3 served as the cut-off.[77] Participants were classified as having neglect if at least one of the tests indicated impairment.

Cognitive assessment was conducted by a neuropsychologist and comprised the Trail Making Test (parts A and B), the Drive Home Maze Test, the Stroop test, the Starry Night Test, the Road Law and Road Craft Test, and the simulator-based Sensorimotor Coordination Test. All neuropsychological tests were administered by the same qualified clinical neuropsychologist.

3.2.2.1. Trail Making Test

The Trail Making Test (TMT) is a commonly used tool in clinical neuropsychology. It is applied on its own or as part of a larger test battery, and is considered useful for examining processing speed, sequencing, visual scanning, and cognitive flexibility.[78] Part A of the test asks the person to draw a line connecting numbers from 1 to 25 in order. Part B is more demanding: numbers and letters have to be connected in an alternating sequence (1–A–2–B, and so on). Completion time is the main outcome in both parts.[79] While Part A mostly reflects visual search and general psychomotor speed, Part B also requires flexible switching between two sequences and is therefore often seen as more sensitive to executive problems. The TMT has been applied in driving research for decades and is consistently reported as one of the strongest predictors of driving fitness.

3.2.2.2. Drive Home Maze Test

The Drive Home Maze Test (DHMT) is part of the Occupational Therapy–Driver Off Road Assessment Battery (OT-DORA), which was designed to screen cognitive and visuomotor abilities relevant for driving before undertaking an on-road test.[80] In this paper-and-pencil task, the examinee uses a pen to trace a path through the maze as quickly as possible, with completion time serving as the main outcome measure. The test taps into planning, decision-making, processing speed, attention, and visuoconstructional skills, and has been shown to predict on-road performance with good reliability. Because it is

easy to administer and does not require special equipment, the DHMT has become a practical and widely used component of driver assessment.[81]

3.2.2.3. The Stroop Color and Word Test

The Stroop Color and Word Test (SCWT) is a widely used tool to measure executive functions, especially inhibition.[82] In the classic version, participants read color words, name color patches, and then have to say the ink color of incongruent words (for example, the word "red" printed in blue ink). The main outcomes are reaction time and number of errors, with the incongruent condition usually being the most difficult.[83] The test not only reflects interference control but also gives information about processing speed, attention, and flexibility, and it has often been used in driving-related studies. In our work, the test was run as a simple Windows desktop application. Participants answered with color-coded keys on the keyboard, while the software measured both response times and mistakes. This made it possible to capture performance with good precision in terms of both speed and accuracy.

3.2.2.4. The Starry Night Test

The Starry Night Test (SNT), developed by Deouell and colleagues, was designed to overcome limitations of conventional paper-and-pencil neglect tasks by combining spatial attention and reaction time measurement.[53] In this task, visual targets briefly appear within a virtual 7×7 grid overlaid on a background of flickering distractors. Participants are asked to respond as quickly as possible when the red target appears. Across 490 trials, both detection accuracy and reaction times are recorded. In our study, we used two outcome variables: the mean response time across the whole display (SNT-mean) and the percentage difference in response time between the left and right sides (SNT-percentage). We obtained permission from the original developers and implemented the test in Unity (C#) as a standalone Windows application.

3.2.2.5. Hungarian version of the Road Law and Road Craft Test

Knowledge of traffic regulations was assessed with the Hungarian version of the Road Law and Road Craft Test (RLRCT-H). The original test consists of 14 items that cover both road law (including signs) and road craft (practical aspects of driving and factors influencing it).[84] It not only tests factual knowledge but also engages problem-solving, sustained attention, and visual perceptual skills, as many items require interpreting

diagrams and pictures. In the Hungarian version, the test was extended with static items and dynamic video clips taken from the official national driver's license examination database, making it directly relevant to national driving conditions.

3.2.2.6. Sensorimotor Coordination Test

The Sensorimotor Coordination Test (SCT), administered via the EM-05.93 Combined Driving Simulator which is designed to evaluate real-time sensorimotor integration relevant to driving. During the task, participants are required to respond to one of five possible visual stimuli presented on a screen, each mapped to a specific motor response (e.g., turning the steering wheel or pressing one of the foot pedals). The device is distributed by Struktura Instruments Co. Ltd. It does not have an English manual, and no standardization study has been published to date, nor has it been reported in international publications. Its use has mainly been limited to Central Europe. For these reasons, the SCT was not included in the published paper, but it is described here in the dissertation together with the results.

Although the EM-05.93 simulator has not been validated in clinical populations, it was included in this study because it has been routinely used in everyday clinical practice at the Rehabilitation Clinic of Semmelweis University. Obtaining empirical data was therefore considered important, as current cut-offs are based on healthy controls rather than on patients with real driving assessments. According to driving assessment literature, deriving cut-off values solely from healthy samples can be misleading, since clinical relevance should be established against on-road outcomes rather than normative performance alone. By including the simulator, the present research aimed to provide additional, practice-based information that is currently absent from the published literature.

3.2.2.7. On-road test

Finally, the on-road test, considered the gold standard in driving research, was carried out within two weeks of the cognitive assessment. This 40-minute standardized course, administered by a licensed driving instructor blinded to previous results, included a wide range of traffic conditions. Performance was rated according to predefined criteria covering rule compliance, hazard perception, vehicle control, and situational awareness.

Based on this evaluation, patients were classified as either fit or unfit to drive, in line with national regulations.

3.2.3. Data analysis

Group-level differences were examined with nonparametric statistics: the Mann–Whitney U test was applied for continuous variables, while categorical variables were compared using chi-square (χ^2) tests. For reaction time–based measures (i.e., simulator based SCT results and the Stroop test), inverse efficiency scores (IES). The inverse efficiency score was applied as a joint measure of reaction speed and accuracy.[85]

Instead of treating response time and errors separately, IES integrates them into a single index by dividing the mean reaction time (RT) by the proportion of correct responses (ACC):

$$IES = \frac{RT}{ACC}$$

This calculation yields an estimate of the average time needed for a correct response. In practice, lower values indicate faster and more accurate performance, while higher values point to slower or more error-prone responding. The advantage of IES is that it adjusts for the typical speed-accuracy trade-off: a participant cannot achieve an artificially good score by responding very quickly at the cost of high error rates.

Originally developed in experimental psychology, IES has increasingly been used in visual-motor tasks and ageing research. It provides a more balanced picture of cognitive performance than reaction times or accuracy alone, and is especially useful when subtle deficits are expected, such as in tasks requiring divided attention, inhibition, or rapid sensorimotor integration.

IES clearly separates participants who respond both quickly and accurately from those who are slow and make frequent errors. It is less informative in profiles where responses are slow but correct, or fast yet inaccurate. In this study, these patterns were not analysed as distinct groups, as safe driving depends on both timely responses and correct decision-making. A response that is accurate but delayed, or rapid but erroneous, can carry a comparable safety risk in real traffic.

For the SNT mean variable, inverse efficiency scores were not computed. If no response was given within the available 3000 ms, it was registered as an omission and counted as a 3000-ms reaction time in the mean calculation, ensuring that missed trials were represented as prolonged response times.

To explore predictive factors of driving performance, logistic regression models were employed. This method was selected because it is well suited to relatively small datasets and offers a high degree of interpretability.[86] Predictor selection was carried out by testing every possible combination of variables in an exhaustive search procedure, beginning with a full model and systematically evaluating all reduced models. General descriptive variables - including both demographic and clinical characteristics that showed associations with driving outcome (e.g., age or the time elapsed since the last driving activity) - were not entered into the model. Preliminary analyses indicated that they would not meaningfully increase predictive accuracy, while their inclusion would raise the number of predictors and thereby increase the risk of overfitting in a small clinical dataset. Model performance was evaluated through leave-one-out cross-validation (LOOCV). In this approach, the dataset is split in such a way that every single observation is used once as a test case while the remaining n-1 cases form the training set. The model is repeatedly trained on n-1 observations and tested on the one that was left out, until all data points have been used in turn. The final performance estimate is the average across all iterations. The strength of LOOCV is that it makes full use of the available data: every observation contributes to both model training and testing. Because each training set is almost the size of the full dataset, the bias of the performance estimate is very low. This property makes LOOCV particularly valuable when working with relatively small datasets, as is often the case in clinical research.

In addition, 95% confidence intervals were derived using bootstrap resampling with 2000 iterations.[87] Feature selection was based on an exhaustive search strategy, systematically testing all possible predictor combinations to identify models with the highest cross-validated ROC-AUC. The final models retained four and five predictors, respectively. To aid interpretation and highlight the relative contribution of each predictor, SHapley Additive exPlanations (SHAP) values were computed.[88] The method is based on the Shapley value from cooperative game theory, where each feature of the dataset is treated as a "player," and the model's prediction is divided among them based on their

contribution. In practice, SHAP assigns each feature a value that reflects how much it added (positively or negatively) to a given prediction. The main advantage of SHAP is that it produces consistent and accurate explanations, meaning that the sum of all feature contributions matches the model's prediction.

All statistical analyses were performed in Python 3.11.5. Data handling and preprocessing relied on pandas (v2.2.2), nonparametric tests were conducted with scipy (v1.13.0), logistic regression and cross-validation procedures were implemented using scikit-learn (v1.4.2), and SHAP values were derived with the shap package (v0.45.0). Throughout the analysis, best practices were followed to reduce the risk of overfitting and to ensure reproducibility.

4. Results

4.1. Narrative review

4.1.1. Analysis of traffic statistics

The rate of fatal and serious injuries shows a typical U-shaped curve in the age distribution, as the youngest and the oldest are the most affected (see Figure 3.).[89,90] These results often form the basis for the notion that chronological age is in itself a risk factor in traffic safety.[64]
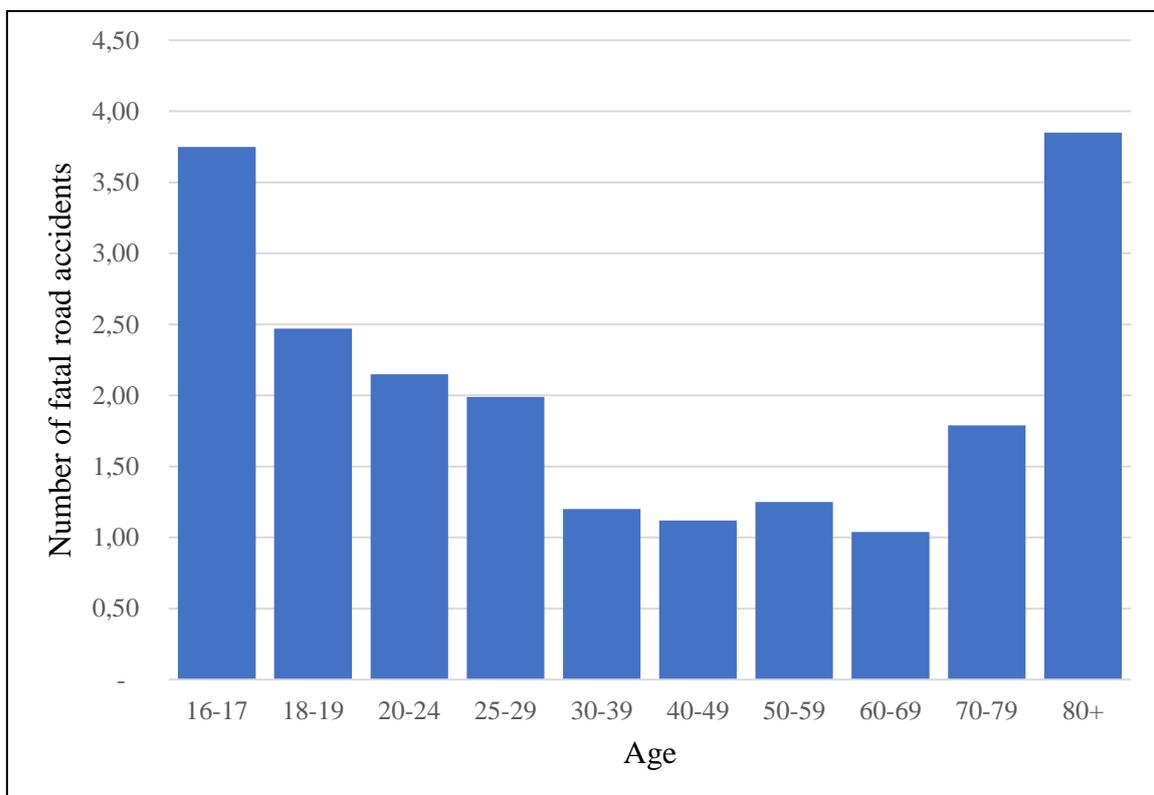
Figure 3. Number of fatal traffic accidents per 100 million kilometres driven.[64,87] Reproduced with permission of the author

However, if we take all traffic accidents regardless of their outcome, the pattern changes (see Figure 4.). The crash rate of the oldest shows a modest rise, but in this diagram the risk of the young drivers is evident.[89] A broadly accepted solution to this paradox lies in the physical vulnerability of the elderly. As this age group is significantly more fragile, they tend to suffer serious injuries in crashes, thus they are overrepresented in traffic

accident reports. The age of the in-fault driver significantly influences the severity of the outcome, as Ichikawa et al.'s review argues.[91]

Compared to younger drivers, accidents caused by the elderly come with fewer personal injuries. At the same time, the in-fault elderly driver gets injured more often compared to young drivers.
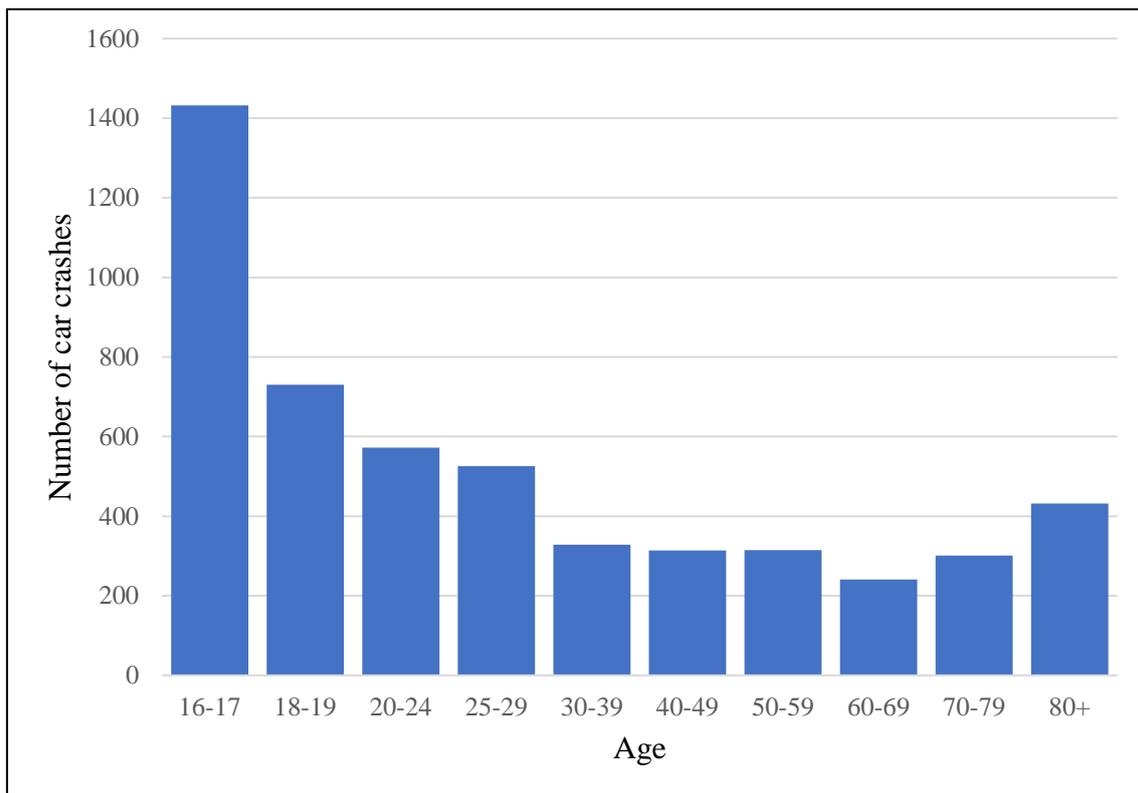
Figure 4. Total number of traffic accidents, independent of outcome severity, per 100 million vehicle kilometres travelled.[64,87]

4.1.2. The effects of age-dependent medical screening

It is evident that with ageing, sensory, perceptual, cognitive, and physical abilities decline, resulting in an elevated crash risk.[15] In order to reduce this risk, many countries have implemented age-dependent assessments of driving abilities. In their 2015 systematic review, Siren and Haustein investigated the effect of these screening programs on traffic statistics.[20] The study reached the surprising conclusion that age-dependent

medical screening does not reduce crash rates and may even have a negative impact on the elderly.

One striking example is Denmark. In 2006, the country tightened the regulations of ageing drivers by implementing two common cognitive tests, the Mini-Mental State Examination (MMSE) and the Clock Drawing Test. In their publication in 2012, Siren and Meng investigated the effect of this change in the screening protocol in the three years before and after the implementation.[92] The car accident statistics were generally worse after the tightening, as there was a 7% increase in car accident fatalities among the young and 14% among the over-70 age group, but the difference between the groups was not significant. However, there was a 38% increase in fatalities in the elderly as pedestrians, cyclists, or moped users. Among the young, there was a 9% increase, and the difference between the groups was significant. The authors concluded that age-dependent medical screening is a typical example of political decisions that make sense intuitively but fail to reach the desired goal.

Another compelling example is the comparison between Sweden and its neighbor, Finland. While Sweden has no mandatory age-dependent medical screening, in Finland, drivers over the age of 70 are required to attend regular medical check-ups.[93] The two countries' car crash statistics did not differ significantly. However, in Finland, there are notably more elderly individuals involved in severe or fatal traffic accidents as pedestrians or cyclists. According to the authors' conclusion, those who lose their driving licenses without a proper indication or thorough assessment are forced to rely on other means of transportation. This can leave them less protected on the road, which is why medical screening might sometimes have a neutral or even negative impact on overall traffic safety statistics.

An Australian study compared the traffic accident statistics of the six states.[94] According to their results, the lowest rates of fatalities and serious injuries were registered in Victoria, the only Australian state with no obligatory medical screening. The presumed explanation for these unfavourable findings is that the methods commonly used to determine fitness to drive lack sufficient scientific support.

4.1.3. Summary of key findings from the review

The literature consistently showed that chronological age by itself is a poor indicator of driving fitness. Studies agree that driving performance is better explained by a person's functional cognitive status, especially in domains such as attention, executive control and visuospatial processing. Several papers reported that some individuals with mild cognitive impairment, and even those in the early stages of dementia, can continue to drive safely for a limited time if properly assessed.

The review also highlighted that age-dependent medical screening programmes offer little benefit and at times may even be counterproductive. This is largely due to the tendency to treat older adults as a homogeneous group, despite wide variation in the pace of physiological ageing. An additional difficulty is that screening often poses the wrong clinical question: the issue is not whether a person is "medically healthy," but whether they can operate a vehicle safely. When this distinction is missed, inappropriate tests are selected and thresholds are applied based on general cognitive norms rather than driving-related risk.

Evidence from international studies further indicated that trichotomous decision methods yield safer classification than binary cut-offs, as they reduce false conclusions in borderline cases. Widely used screening batteries such as OT-DORA have been shown to perform inconsistently when interpreted without such a structured decision framework.

4.2. Original study

4.2.1. Clinical and demographic variables

Of the 115 participants included in the analysis, 80 (70%) were classified as fit to drive and 35 as unfit, according to the outcomes of the on-road evaluation. A detailed overview of the demographic and clinical characteristics of the two groups is provided in Table 1. Interestingly, the fit-to-drive group still included 12 patients with clinically diagnosed hemispatial neglect at the time of enrolment, out of 26 neglect cases in total. Within this group, 9 of 22 patients with left-sided neglect passed the on-road assessment, as did 3 of 4 patients with right-sided neglect.

Table 1. Mann-Whitney U tests were used for: age, sex, driving experience, time since last driven, time since stroke; Chi-square tests were used for affected hemisphere and hemispatial neglect.[73]

Variable	All n = 115	Passed n = 81	Failed n = 34	p-value
Age, years, mean ± SD	53 ± 8	52 ± 8	56 ± 7	**0.014**
Sex, % female	22.6	22.2	23.5	1.0
Affected hemisphere, % left	38.3	40.7	32.4	0.465
Driving experience years, mean ± SD	33 ± 9	30.9 ± 9	31.4 ± 9.3	0.675
Time since last driven, days mean ± SD	118 ± 253	150 ± 268	240 ± 208	**0.000**
Time interval between stroke and on-road test, days mean ± SD	553 ± 1140	650 ± 1352	330 ± 278	0.189
Patients with the clinical diagnosis of hemispatial neglect (%)	26 (22.6)	12 (14.8)	14 (41.2)	**0.007**

4.2.2. Logistic regression model

While several individual predictors showed strong associations with driving fitness, logistic regression provided the best overall predictive performance (see Figure 5.). The highest cross-validated ROC-AUC values were achieved with the following optimized feature sets:

- Optimal 3 features: Stroop, RLRCT-H, and SNT-mean
- Optimal 5 features: DHMT, RLRCT-H, SNT-percentage, SNT-mean, and Stroop
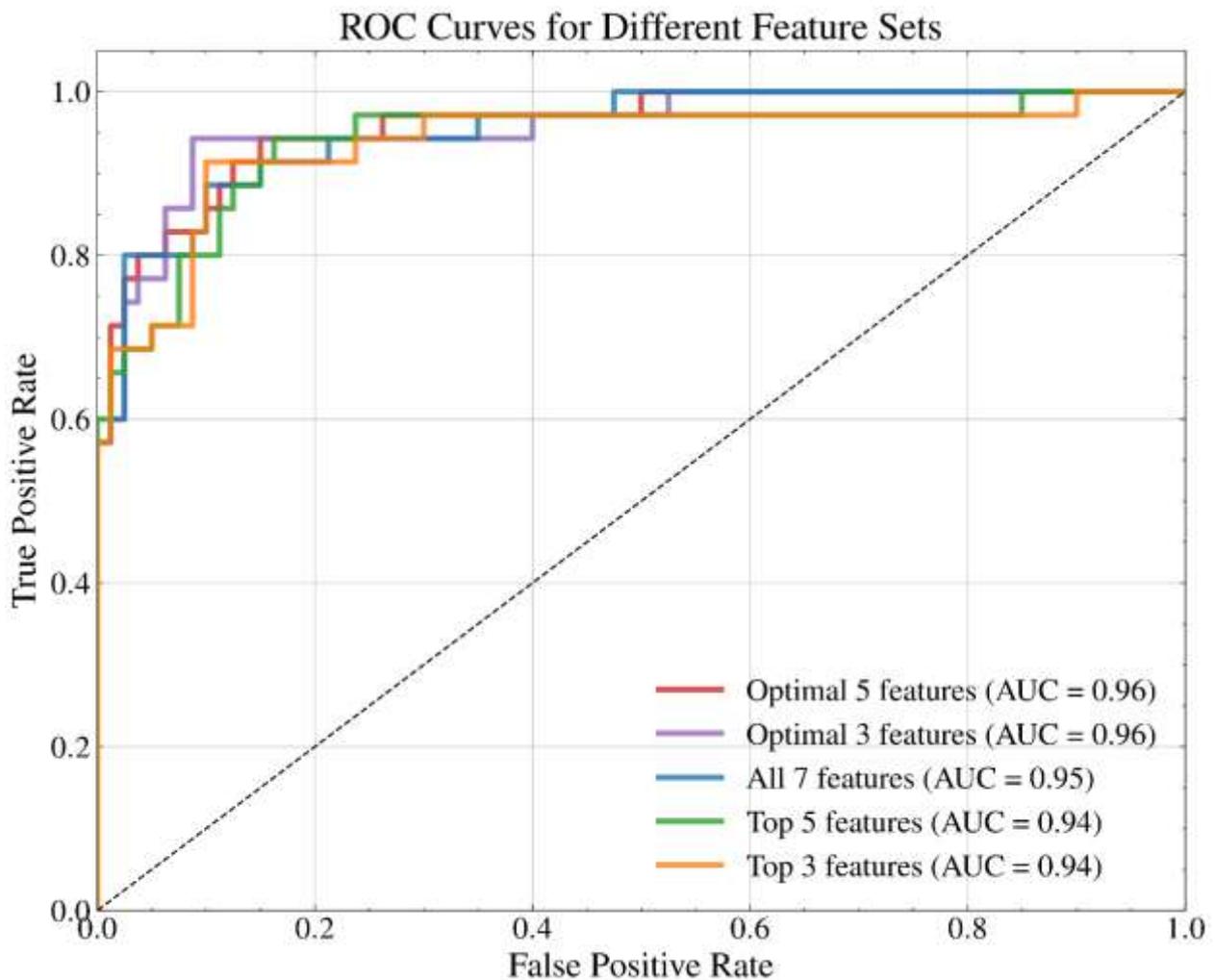
Figure 5. Predictive validity of individual features and logistic regression models trained on various feature sets.[73] ROC-AUC (Receiver Operating Characteristic - Area Under the Curve) measures a model's ability to distinguish between "safe to drive" and "unsafe to drive" patients across different decision thresholds. A higher ROC-AUC indicates better overall performance, with a value of 1.0 representing perfect discrimination and 0.5 indicating random guessing.

4.2.3. Multicollinearity among variables

A correlation analysis was performed to evaluate potential multicollinearity among the predictor variables. Excessively high collinearity is typically indicated by coefficients approaching or exceeding $r = .80$, as such values can distort regression estimates and diminish the interpretability of individual predictors.[72,95] Although several associations reached statistical significance in the Pearson analysis, this pattern was expected given shared cognitive demands across the tasks (see Table 2.). None of the coefficients approached the $r = .80$ threshold, indicating that the observed relationships were not large enough to compromise the predictive model.

Moderate associations were mainly observed among executive and processing-speed measures: for instance, TMT-B correlated with Stroop performance ($r \approx .65$), and both showed negative relationships with Traffic Knowledge ($r \approx -.52$ to $-.68$). Reaction time on the Starry Night Test also correlated with visuomotor speed (TMT-A; $r \approx .43$), suggesting shared processing requirements without redundancy. These findings align with partially overlapping yet distinct cognitive operations, supporting the inclusion of all predictors in the multivariate model.

Table 2. Correlation matrix of cognitive and driving-related variables
Pearson correlation coefficients describe associations among the variables included in the model. Asterisks indicate significance levels (*p < .05, **p < .01). No coefficient exceeded r = .80, indicating relatively low level of multicollinearity. Moderate relationships were mainly observed between executive and processing-speed measures (e.g., TMT-B and Stroop).

Variable	DHMT	TMT-A	TMT-B	RLRCT-H	SNT percentage	SNT mean	Stroop
DHMT	-						
TMT-A	0.350**	-					
TMT-B	0.516**	0.595**	-				
RLRCT-H	-0.435**	-0.518**	-0.673**	-			
SNT percentage	-0.036	0.293**	0.165	-0.158	-		
6. SNT mean	0.299**	0.433**	0.320**	-0.400**	0.519**	-	
7. Stroop	0.547**	0.573**	0.655**	-0.681**	0.082	0.222*	-

4.2.4. SHapley Additive exPlanations analysis

The repeated inclusion of Stroop, RLRCT-H, and SNT-mean across both models underlines their particular relevance. SHAP analysis confirmed this pattern, identifying Stroop as the single most influential factor, followed by RLRCT-H and SNT-mean. Elevated Stroop and SNT-mean scores, together with lower RLRCT-H performance, were consistently linked to a higher probability of being classified as unfit to drive (see: Figure 6.).
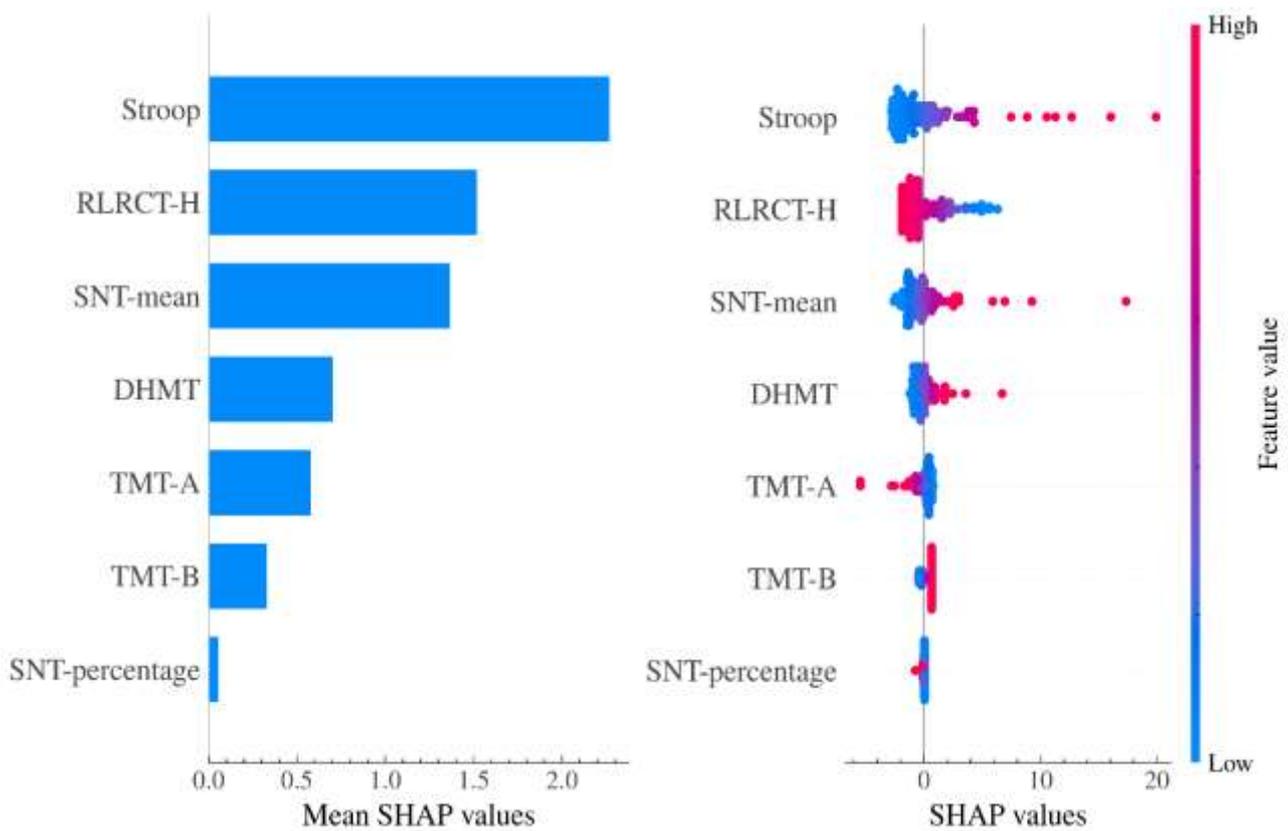
Figure 6. Mean SHAP values and feature contributions.[73]
The left plot displays the average impact of each feature on model predictions, while the right plot illustrates how specific feature values contribute to predictions.
In the right plot, each dot represents a patient, with red indicating high feature values and blue indicating low values. The x-axis shows the effect of each feature on the prediction: dots positioned further to the right indicate a stronger contribution to predicting 'unsafe to drive,' while those on the left suggest a contribution toward 'safe to drive.'

4.2.5. Trichotomous model

Inspection of the predicted probability distributions suggested that 85% of participants could be clearly allocated into either the fit or unfit category. For the remaining 15%, prediction confidence was low, leading to the creation of an intermediate "undecidable" category (see Figure 7.).
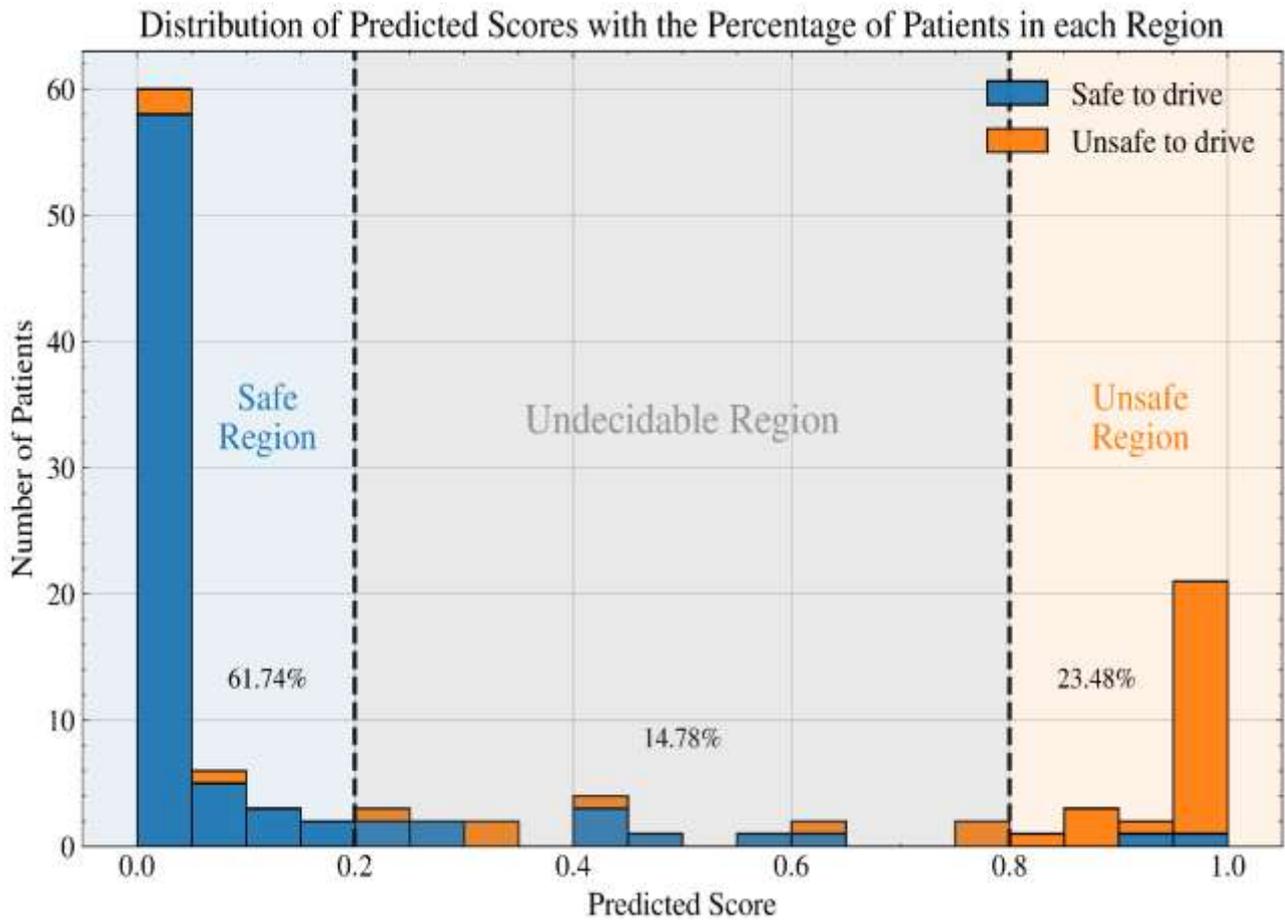
Figure 7. Trichotomization of predictions.[73] It can be observed that 85% of patients were clearly categorized as either safe to drive or unsafe to drive. The remaining 15% fall into an undecidable region, where further testing is recommended to better assess their driving ability.

Introducing this trichotomous classification improved model robustness and reduced the likelihood of misclassification in borderline cases. Although the confidence intervals remain wide due to the modest sample size, the observed tendencies point toward meaningful gains in predictive accuracy (see Table 3.).

Table 3. Performance of the model with and without trichotomization. Values in parentheses represent 95% confidence intervals.[73]

	Percentage of classified patients	ROC-AUC	Specificity	Sensitivity	Precision	Accuracy
Without trichotomization	1	0.96 (0.91 - 0.98)	0.95 (0.91 - 0.99)	0.84 (0.71 - 0.95)	0.89 (0.78 - 0.97)	0.92 (0.86 - 0.98)
With trichotomization	0.87 (0.80 - 0.98)	0.97 (0.92 - 1.00)	0.98 (0.95 - 1.00)	0.88 (0.77 - 0.97)	0.95 (0.88 - 1.00)	0.95 (0.91 - 0.99)

Notably, in the initial analyses the SCT variable appeared in both the Optimal 3 and Optimal 5 feature sets, and SHAP identified it as the strongest single contributor to the model. Since the SCT has not yet been externally validated, it was excluded from the final models. Paradoxically, removing the best-performing variable did not change the key performance measures of the model, as the model adapted by adjusting its classification boundaries. The main effect of this adjustment was that more patients were placed into the inconclusive category: with the SCT included, definite classifications were made for 90.4% of cases, whereas after its removal this rate dropped to 85.2%.

5. Discussion

5.1. Older drivers

Accident statistics indicate that older drivers pose a slightly increased safety risk on the road.[21] Nevertheless, chronological age by itself proved to be a weak predictor of driving ability, indicating that the elderly do not constitute a homogeneous group.[20] According to traffic statistics, more complex traffic situations placing demand on sustained attention and requiring rapid reactions are particularly problematic even for healthy older adults. At the same time, a considerable number of individuals with MCI or even dementia may remain capable of safe driving for a limited period.[23] These observations highlight the fundamental principle in rehabilitation that diagnostic labels are less relevant than the individual's actual functional state.

In many countries, periodic health screenings of older drivers are standard practice. Yet evaluations of their effectiveness often reach a surprising conclusion: such screenings do not reliably improve accident statistics.[20] This does not mean that medical review is unnecessary, but it does draw attention to the limited value of the current methods in identifying truly high-risk drivers. In Hungary, as noted by Henézi, Horváth, and Szegedi, there is no unified protocol for assessing driving fitness in the elderly.[96] As a result, decisions can be subjective, and most of the responsibility falls on the general practitioner. However, GPs do not have simple and well-validated cognitive tools that could support these judgments within the limited time and resources of primary care. While they can refer cases to the official traffic psychology examination (PÁV), its capacity would be overwhelmed if tens of thousands of elderly drivers required annual testing. In practice, therefore, most decisions must be made at the GP level.

To support this task, the introduction of a brief neuropsychological screening battery has been proposed. Such a tool should take no longer than 10–15 minutes, rely on validated measures, and be usable even by trained staff without specialized qualifications. A combination of the Montreal Cognitive Assessment (MoCA), the Trail Making Test, and a simple computer-based reaction time task, interpreted within a trichotomous framework, could provide a reasonable basis for decision-making. The principle can be illustrated with the Trail Making Test Part B, a five-minute measure of cognitive flexibility and set-shifting. Gibbons and colleagues, studying 83 participants, showed that

those completing the test in under 80 seconds were consistently classified as fit on-road, while those needing more than 178 seconds were uniformly unfit.[66] Scores between these thresholds indicate the need for further evaluation. Models based on this logic offer a promising framework, but must first be validated in Hungarian samples before widespread use. Importantly, such screening has an added value, since domains relevant to driving overlap substantially with traditional dementia screening. Earlier identification of mild cognitive impairment or dementia could therefore also be facilitated, enabling earlier treatment.

If screening results are inconclusive, referral to further traffic psychological evaluation (PÁV) is warranted. In cases of neurological disease such as stroke or dementia, more detailed neuropsychological assessment should be considered. This is because several cognitive disturbances that strongly affect driving, such as mild or chronic neglect, visual extinction, or perseveration and shifting deficits, may not be adequately captured by standard traffic psychology tests.

Finally, on-road assessment under real traffic conditions should be introduced in Hungary. Unlike in several other countries, such testing is not currently available, primarily due to high costs. Nevertheless, the literature suggests that even comprehensive, evidence-based medical and psychological tests leave a group of individuals whose driving ability can only be judged reliably through on-road evaluation.[58,97,98] Implementing such procedures would require closer collaboration between health and traffic professionals and could improve both the objectivity and the safety of decision-making.

5.2. Novel model for driving after stroke

To our knowledge, this is the first Hungarian study that systematically examined post-stroke driving fitness with a standardized on-road test as the main outcome.[75] In our cohort, around 70% of patients were classified as fit to drive, which is in line with international results. This finding confirms that the majority of stroke survivors retain sufficient ability to pass an on-road test, although careful assessment remains essential. The novel predictive model performed with strong discriminative accuracy. By incorporating a third, indeterminate outcome category, the risk of misclassification was reduced and clinical decision-making became more reliable. The ROC-AUC of 0.95 is

comparable to the results of a recent prospective study applying a neural network approach in a Japanese cohort of 55 stroke survivors.

5.2.1. Trichotomization

Based on the methodological considerations arising from the narrative review, there is a clear need for assessment approaches that move beyond strictly dichotomous judgments and are able to accommodate uncertainty in a transparent and systematic manner.[64,99] In this regard, trichotomous decision-making approaches described in the international literature offer a conceptually coherent alternative, especially when adapted to national clinical and legal contexts.

The application of a trichotomous decision approach in the present dissertation follows directly from these methodological considerations. International findings consistently indicate that binary cut-off–based decisions are ill-suited to reflect the continuous and overlapping nature of cognitive abilities relevant to driving.[100,101]

At the same time, the Hungarian literature provides little empirically grounded guidance on how diagnostic uncertainty should be handled in fitness-to-drive decisions.[96] Within this context, trichotomization can be understood not merely as a technical modification of classification, but as a conceptual response to an identified methodological gap. By explicitly introducing an indeterminate category, the proposed model treats uncertainty as an inherent element of clinical decision-making rather than as a source of error.[32]

Accordingly, the trichotomous model presented in this dissertation builds on the theoretical and methodological conclusions of the narrative review and translates them into a framework that can be a subject of empirical examination.

The interest in trichotomization emerged as a response to the persistent misclassification problem in binary prediction of driving fitness.[58] Traditional cut-off based approaches force a yes/no decision even in borderline cognitive profiles, which inevitably leads to a high rate of both false positives and false negatives. To address this issue, several authors proposed classifying individuals into three categories: safe, unsafe, and indeterminate. Early applications of trichotomization yielded promising results. Kay et al. reported that introducing a third category significantly improved specificity and sensitivity compared to traditional, binary classification.[102] Similarly, Dobbs and Schopflocher demonstrated that trichotomization enabled more accurate risk stratification among medically at-risk

drivers, although at the cost of a very large proportion of indeterminate outcomes (49%).[103] To mitigate this limitation, serial trichotomization was introduced by Gibbons et al. in 2017.[66] Instead of relying on a single test with two thresholds, their model sequentially applied multiple cognitive measures in a decision funnel. This approach reduced the proportion of inconclusive classifications to 21.7%, demonstrating that iterative decision-making improves classification efficiency without a full dependence on a single test score.

However, despite these advances, most serial trichotomization studies share methodological limitations. First, cut-off values were derived by forcing 100% sensitivity and 100% specificity within the same dataset, a retrospective optimisation that is highly prone to overfitting. As noted by Hoggarth et al., models optimised in-sample tend to overestimate their true predictive accuracy and rarely generalise beyond the development sample.[72] Second, earlier models offered no form of validation, making it impossible to judge the stability of the reported cut-offs. Third, nearly all trichotomization research so far has been conducted in heterogeneous geriatric samples or mixed neurological populations, whereas driving after stroke represents a distinct clinical context with different cognitive risk patterns. As Molnar also emphasised, the clinical utility of a screening method heavily depends on the population for which it was developed.[58]

Our study differs from classical serial trichotomization in three key methodological aspects. First, instead of forcing extreme thresholds, cut-off values were derived using statistical optimisation based on ROC characteristics, which is less sensitive to outliers and avoids exaggerated classification performance.[104] Second, we embedded trichotomization into a probabilistic modelling framework by using a trichotomous logistic regression model. Consequently, classification does not result from arbitrary thresholds but from probability zones that reflect clinically meaningful risk levels. Third, to address the overfitting problem highlighted by Hoggarth et al. (2015), we used leave-one-out cross-validation (LOOCV), a method that provides a realistic estimate of model performance by repeatedly testing predictions on cases not used for model training.[105]

A further methodological consideration concerns the role of test selection in trichotomous classification. In our initial analyses, the SCT variable emerged as a highly influential predictor: it appeared in both the Optimal 3 and Optimal 5 feature configurations, and

SHAP analysis consistently ranked it as the strongest single contributor to model predictions. However, despite its statistical utility, the SCT was excluded from the final models because of the lacking external validation. Interestingly, removing its best-performing predictor did not lower the overall discriminative performance of the model, but it changed the way our model handled diagnostic uncertainty. Instead of reducing accuracy, the model adapted by becoming more cautious: it reduced the number of definitive classifications and assigned more patients to the inconclusive category. With the SCT included, 90.4% of cases received a clear prediction, whereas after its removal this proportion decreased to 85%, indicating a deliberate shift toward caution rather than error. This pattern highlights an important practical insight: model optimization is not solely a statistical exercise but depends equally on the quality of the cognitive tests used. High-quality, validated predictors allow classification models to make confident decisions more often, whereas less informative or inconsistent tests force a larger reliance on the uncertainty zone. Thus, improving off-road driver assessment is not only a matter of advanced modelling but also of refining the test battery itself.

In practical terms, our refined trichotomization approach aims to deliver reliable classification only when the statistical evidence is clear, while patients with intermediate probability are directed to on-road assessment. This selective referral strategy has been recommended previously to reduce unnecessary on-road evaluations, lower cost, and minimise traffic exposure in unsafe drivers.[66] Unlike earlier models developed on mixed samples, our model was trained specifically in a stroke population, where classification uncertainty is particularly high due to heterogeneous lesion profiles and frequent co-occurrence of visuospatial deficits.[65]

Overall, our results show that trichotomization remains a valuable decision framework, but its clinical utility depends heavily on how it is implemented. The transition from fixed thresholding to probability-based classification with built-in validation represents a methodological advancement that improves both robustness and clinical interpretability. In the following section, we discuss how the cognitive predictors included in our model relate to driving performance.

5.2.2. Cognitive functioning and fitness to drive

Driving is widely recognised as a high-level behavioural activity that relies on the interplay of several cognitive systems rather than any single function. Earlier reviews reached similar conclusions, noting that driving draws simultaneously on visual monitoring, attention, decision-making, and motor planning.[35,106] Our findings are consistent with this multidomain view. Instead of one dominant neuropsychological predictor, the literature describes driving competence as a cognitive profile, formed by complementary sources of information.

Within this profile, processing speed emerged as a central component. This result was expected, given that real traffic demands continuous, real-time reactions to changing input.[107] Drivers who process information too slowly miss opportunities to correct errors or avoid hazards - an observation reported repeatedly in ageing and stroke populations.[41] In our study, processing speed was primarily captured by reaction-time components of the Trail Making Test (TMT) and Stroop performance. Interestingly, the inverse efficiency index provided more predictive value than raw reaction times, which suggests that speed alone is insufficient; what matters in driving is how efficiently a person balances speed and accuracy under cognitive load.

A second strong contributor was executive functioning, particularly the ability to flexibly shift between tasks and suppress irrelevant responses.[78] This aligns with previous reports linking executive problems to tactical driving errors such as poor lane planning, slow error correction, or rigid behaviour in traffic.[108,109] In our model, this domain was indexed by both TMT-B and Stroop interference scores. The Drive Home Maze Test (DHMT) also reflects this domain, likely because it requires rule maintenance and stepwise planning rather than pure visuoconstructional skill.

The role of attention, especially divided attention, also deserves emphasis. In daily driving, monitoring mirrors, signage and surrounding vehicles is not a sequential process but a simultaneous one.[110] Prior findings have shown divided attention to be one of the strongest predictors of on-road difficulties after stroke.[36,111] Our results are consistent with this but add a nuance: rather than static attention tasks, dynamic performance-based measures, such as the Sensorimotor Coordination Test (SCT), were more informative. They more closely approximate the temporal demands of traffic. Although the SCT was

excluded from the final model for methodological reasons, its early impact reinforces the idea that motor–cognitive coupling is an integral part of driving ability.

Visuospatial processing also contributed significantly. The literature often links visuospatial impairment to lane drift, difficulty estimating distance, or misjudging vehicle position.[36,112] The SNT appears to capture this dimension effectively, even in patients without clinically diagnosed neglect.[53] Its dynamic format may explain why: unlike static cancellation tasks, it requires continuous scanning and rapid perceptual judgement.

In addition to these cognitive-perceptual predictors, knowledge of traffic rules and driving logic also added independent value in our model. This is consistent with evidence from SDSA-based studies, suggesting that safe driving depends not only on perceptual speed and attention but also on semantic understanding of traffic conventions.[113] The RLRCT-H served as a compact indicator of this decision-rule competence.

An important methodological point is that we deliberately selected predictors with low redundancy. Hoggarth et al. warned that many cognitive models of driving suffer from high intercorrelation between variables, which weakens validity and exaggerates apparent model strength.[72] In contrast, each predictor in our final model represented a distinct cognitive function, reducing collinearity and supporting more stable classification performance.

Driving-related cognition does not operate in a vacuum.[35] In real traffic, individuals do not rely on isolated cognitive functions but on habits, routines and strategies developed over years of practice. This explains why performance in standardised office-based assessments do not always mirror behaviour in complex, everyday situations. Many stroke survivors, for instance, find ways to work around subtle deficits: they reduce their speed, increase head movements, drive only in familiar areas or avoid peak traffic.[114] These strategies may temporarily compensate for cognitive weaknesses and allow a person to pass an on-road test, despite measurable deficits in processing speed or spatial attention. Because of this gap between controlled test performance and behaviour in naturalistic contexts, it is unlikely that any cognitive model will reach perfect predictive accuracy.

A more realistic approach accepts a degree of diagnostic uncertainty. Our results support this direction by demonstrating that a trichotomous system can manage it productively. Cases with a clear cognitive risk profile can be identified with high confidence, while borderline cases are explicitly flagged for further evaluation rather than misclassified. In that sense, trichotomization is less about statistical elegance and more about responsible clinical decision-making.

Overall, our findings reinforce a profile-based interpretation of fitness to drive. Different cognitive domains do not contribute equally, but each adds a piece of information that forms a coherent risk picture, taken together. This conclusion naturally leads to one domain that deserves closer attention: visuospatial-attentional control. Hemispatial neglect, in particular, is widely cited as one of the strongest cognitive predictors of unsafe driving after stroke. The pattern in our data confirms its importance, but the relationship is more layered than simple deficit–risk models suggest.

5.2.3. Hemispatial neglect syndrome and driving

Hemispatial neglect is one of the most frequently cited cognitive risk factors for unsafe driving after stroke, and in many clinical contexts it is treated almost automatically as a contraindication for returning to driving.[115,116] However, while the functional consequences of neglect are well documented, the assumption that it should serve as a universal exclusion criterion has been based more on theoretical reasoning than on systematic empirical evidence.[117] In contrast, visual field defects such as homonymous hemianopia have been studied in driving research in significantly more depth: several studies indicate that a proportion of individuals with hemianopia can drive safely with compensatory strategies.[118,119] Whether a similar compensatory mechanism may operate in neglect has remained an open question.

A fundamental problem is that neglect is not equivalent to visual field loss. It is not a sensory deficit but a higher-order attentional disorder that affects spatial awareness across multiple modalities.[120] Moreover, neglect is a heterogeneous syndrome: it may manifest in egocentric or allocentric reference frames, may be restricted to personal or extrapersonal space, and may commonly be accompanied by extinction or anosognosia.[46] No single test captures all these aspects, and bedside diagnosis varies widely depending

on the tools used.[121] This complexity may explain why earlier studies treated neglect as a categorical risk without investigating which aspects of neglect matter most for driving.

Our study did not aim to provide an exhaustive assessment of all neglect dimensions. Instead, we selected measures that are commonly used in clinical diagnosis (Bells Test, line bisection) and added a dynamic, computer-based measure of lateralised attention, the Starry Night Test (SNT).[53] The SNT was included because traditional paper–pencil tasks may underestimate neglect in structured environments, whereas time-pressured computerised tasks can detect subtle spatial bias and extinction that emerge under attentional load.[51,122]

Our findings revealed a more nuanced relationship between neglect and driving than expected. Patients diagnosed with neglect using conventional tests were not universally unfit to drive: in our sample, 45% of all neglect patients passed the on-road assessment. When broken down by laterality, the pass rate was 41% in left neglect (right hemisphere stroke) and 75% in right neglect. This laterality effect aligns with previous observations that right hemisphere damage tends to produce more severe and functionally disruptive neglect, whereas neglect after left hemisphere stroke is often milder and more amenable to compensation.[123,124]

These results carry two opposing interpretations, both with important clinical consequences. One possibility is that some neglect patients are indeed able to compensate functionally, at least in structured driving situations, and should therefore not be categorically excluded; instead, individualised on-road assessment would be justified. The alternative interpretation is that passing an on-road test does not guarantee long-term driving safety in neglect, as compensation may only be temporarily sustained during a short evaluation. In that case, our findings would call into question the reliability of current on-road assessments for this group. Our data alone do not resolve which of these interpretations is correct; they point instead to the need for longitudinal follow-up studies with real-world crash risk data.

One of the more unexpected findings in our analysis was related to how the Starry Night Test contributed to prediction. We originally assumed that the lateralised asymmetry index - the difference in reaction time between the two hemifields - would capture spatial bias and therefore be a strong predictor. According to our result, it carried relatively little

weight in the model, while the overall mean reaction time across the whole screen was a superior variable.

This result fits increasingly well with current views of chronic neglect. Several authors have noted that patients can learn to compensate strategically by directing their gaze more frequently toward the neglected side.[125,126] When this happens, the left–right difference naturally shrinks, even though the underlying attention problem has not resolved. The compensation, however, has a price: patients become globally slower, because every target search demands more effort and repeated reorienting. In hindsight, the SNT reaction time variable captured this phenomenon, which is likely why it separated safe and unsafe drivers better than raw asymmetry scores.

These observations also explain why standard cancellation tests often underestimate real functional impact. Paper–pencil tasks are static, predictable and give patients time to self-correct.[127] They rarely reveal the time pressure or dual-task demands of real-life behaviour. In contrast, dynamic computer-based tasks such as the SNT can expose subtle residual neglect and extinction effects that structured tasks fail to detect.[52] In our sample, several patients who appeared "normal" on conventional bedside tests still showed slowed SNT responses - and some of them later failed the on-road assessment. This supports the argument that reaction-based digital assessment should have a stronger role in driving evaluation after stroke, especially when neglect is suspected.

5.2.4. Limitations

Our study has some limitations that should be considered when interpreting the findings. First, data were collected from a single center. Although patients were referred from multiple regions, this setting may still limit how broadly the results can be applied. To improve generalizability, future research should involve multicenter cohorts that reflect a wider range of clinical and demographic backgrounds. Second, while the sample size was adequate for building and validating a logistic regression model, it was not large enough to explore more advanced machine learning approaches such as Random Forests. Given the encouraging performance of our method, future studies should consider larger-scale, and possibly international, collaborations to enable more complex modeling techniques and further strengthen the evidence.

5.3. Implementation to national practice

Fitness-to-drive decisions are demanding in clinical practice because they involve medical risk and legal responsibility, while at the same time influencing a person's autonomy.[14] After stroke, as well as in older age, giving up driving often leads to loss of independence and social withdrawal, which explains why neither overly permissive nor overly restrictive decisions are acceptable.[128] Cognitive screening can support decision-making, but only if results are interpreted within a clear and transparent decision framework.

As in many European countries, there is currently no outcome-based national guideline in Hungary for evaluating fitness to drive.[96] This reflects a methodological limitation rather than a clinical one: until now, no Hungarian study has linked cognitive test results to actual on-road performance, which means that commonly used thresholds are derived from general cognitive norms rather than driving-specific risk levels. The present results provide an initial empirical step toward outcome-anchored decision-making by showing that cognitive performance can be translated into probability-based risk estimates.

Instead of relying on fixed cut-offs or categorical classification, the proposed model introduces trichotomous decision zones. Rather than assuming that every case can be classified as either fit or unfit based on cognitive data alone, the model distinguishes three groups: likely safe, likely unsafe, and uncertain cases. The third category is particularly important, because it redirects ambiguous cases to further assessment instead of forcing premature decisions. In this sense, the model manages uncertainty explicitly, which increases both fairness and traffic safety.

Provided that all relevant test results are available, the model can already be used as a probability-based aid in clinical decision-making. Its utility therefore depends on administering the full predictor set described in this study, as the model estimates risk from the joint contribution of these variables, and a missing value would prevent reliable probability calculation.

Although this model was developed using stroke patients, the decision logic is transferable to other clinical populations, including older drivers, once predictors are adapted to the relevant cognitive profile. Our narrative review highlighted that widely

used tools such as the SDSA or OT-DORA are not inherently flawed but tend to misclassify drivers when used in isolation.[64] The main challenge is therefore not selecting the "perfect" cognitive test, but developing a consistent way to interpret cognitive information in context.

A three-stage screening pathway appears realistic for national implementation:

1. Initial cognitive pre-screening in general practice
   A brief screening tool would allow general practitioners to identify clear low-risk and high-risk profiles. The trichotomous approach is well suited to this stage, since decisions often need to be made with limited time and partial information.

2. Specialist evaluation for inconclusive cognitive profiles
   Only patients in the uncertain group would proceed to targeted specialist assessment. This reduces unnecessary referrals and concentrates resources on those who most likely need further investigation.

3. On-road assessment limited to unresolved or high-risk cases
   Practical driving tests retain their role but would be used more selectively. This approach not only reduces waiting times but also lowers public cost; in line with current Hungarian practice, the cost of the on-road test could reasonably remain patient-funded when further examination is requested.

Such a staged system would require no major structural change in Hungary. Instead, it would integrate existing medical and administrative elements into a more transparent and reproducible workflow. Introducing this approach into routine practice would first require a few essential steps. The model needs confirmation in larger and more mixed patient groups, and probability thresholds should be tailored to age-related changes if it is to be used with older drivers. Importantly, cognitive findings should never be treated in isolation, but combined with medical status and functional capacity when a licensing decision is made.

6. Conclusion

The aim of this doctoral work was to address several challenges in the field of post-stroke driving assessment.

First, we introduced a new method of data analysis. By combining logistic regression with leave-one-out cross-validation and trichotomization, we demonstrated that it is possible to reduce the risk of misclassification. This approach provided a cautious and clinically useful framework, where uncertain cases were not forced into a binary decision but flagged for further evaluation.

Second, we investigated the role of hemispatial neglect in driving performance. Although neglect is widely assumed to be incompatible with safe driving, our findings showed a more complex picture. A considerable proportion of patients with clinically diagnosed neglect still passed the on-road test, suggesting that compensation is possible in some cases. At the same time, neglect remained a significant risk factor, underlining the importance of its careful assessment.

Third, we evaluated the utility of traditional neglect tests. Tools such as the Line Bisection Test and Bells Test showed only limited ability to predict the outcome of the on-road test. These results highlight the restricted utility of classical paper-and-pencil tools for assessing driving fitness, as suggested by previous research.

Fourth, we assessed the potential of the Starry Night Test. This dynamic, computer-based method captured attentional slowing and asymmetry more sensitively than conventional paper-and-pencil tools. Although other variables such as the Stroop test or the RLRCT-H contributed more strongly to our model, the Starry Night Test still provided significant impact and showed potential as part of a broader assessment battery, even in cases of subtle or chronic neglect.

Fifth, we considered the Hungarian context. In the absence of standardized cognitive protocols for driving assessment, this dissertation summarized international findings and formulated recommendations that could be adapted to national practice. These included the need for validated, short screening instruments and a stepwise process involving general practitioners, neuropsychologists, and on-road testing.

Finally, we explored wider implementation. The principles tested in stroke patients may be extended to other populations, including older drivers. A short, trichotomous screening tool could serve as a practical first-line filter in primary care, allowing uncertain cases to be referred for further detailed evaluation.

7. Summary

This doctoral work investigated the driving abilities of stroke survivors and the elderly, focusing on the possibilities of driving assessments. As physical impairments are more visible and easier to evaluate, more research has focused on cognitive fitness to drive, since even subtle neurocognitive deficits can render a person unfit to drive. In recent decades, several studies have attempted to predict driving fitness using neuropsychological tests measuring processing speed, visuospatial skills, or executive functioning. Evidence showed a significant correlation between specific cognitive test scores and the outcome of on-road assessments, which are considered the gold-standard outcome measure in driving research. Despite these advancements, the rate of false classifications has remained relatively high, as the field continues to struggle with methodological issues such as lack of model validation and the use of dichotomous (i.e., either fit or unfit to drive) classification instead of trichotomization. In our study, we integrated recommendations from the literature into a validated logistic regression model that allows clinicians to abstain from making unfounded decisions when certainty is low.

In a cohort of 115 stroke patients, comprehensive neuropsychological assessment was conducted, followed by a single-blinded, standardised on-road test. Results showed that cognitive testing predicted on-road performance with high precision (ROC-AUC: 0.95). However, this strong psychometric performance was achieved by withholding classification in 15% of the sample. In these patients, cognitive test results were borderline and the risk of misclassification was high, meaning that no reliable decision could be made. This trade-off supports a more cautious clinical practice, based on the idea that in some cases, no decision is better than a wrong one.

Although our model was validated in a stroke population, it might also be adapted for use in assessing older drivers on a larger scale. Previous research has shown that using cognitive tests not specifically designed for driving can actually worsen traffic safety statistics. Future studies are needed to implement a validated, trichotomized model in the form of a brief test battery that general practitioners could use as a first-line tool when evaluating older drivers.

8. References

1. Hwang S, Song CS. Driving Rehabilitation for Stroke Patients: A Systematic Review with Meta-Analysis. Healthcare. 2023 Jun 2;11(11):1637.

2. Fisk GD, Owsley C, Pulley LV. Driving after stroke: Driving exposure, advice, and evaluations. Arch Phys Med Rehabil. 1997 Dec;78(12):1338–45.

3. Mauri M, Sinforiani E, Cuzzoni MG, Bono G, Zucchella C. Driving habits in patients with dementia: a report from Alzheimer's disease assessment units in northern Italy. Funct Neurol. 2014;29(2):107–12.

4. Rosen HE, Bari I, Paichadze N, Peden M, Khayesi M, Monclús J, Hyder AA. Global road safety 2010–18: An analysis of Global Status Reports. Injury. 2025 Jun;56(6):110266.

5. Hungarian Central Statistical Office. 4.2.2.1. Személysérüléses közúti közlekedési balesetek száma negyedévenként, vármegye és régió szerint [Internet]. Available from: https://www.ksh.hu/stadat_files/ege/hu/ege0077.html

6. United Nations Road Safety Collaboration - World Health Organization. Global plan for the decade of action for road safety 2021–2030 [Internet]. 2020. Available from: https://www.who.int/publications/m/item/global-plan-for-the-decade-of-action-for-road-safety-2021-2030

7. Singh H, Kushwaha V, Agarwal AD, Sandhu SS. Fatal Road Traffic Accidents: Causes and Factors Responsible. J Indian Acad Forensic Med. 2016 Mar;38(1):52–4.

8. Visser E, Pijl YJ, Stolk RP, Neeleman J, Rosmalen JGM. Accident proneness, does it exist? A review and meta-analysis. Accid Anal Prev. 2007 May;39(3):556–64.

9. Wilde GJS. Risk homeostasis theory: an overview. Inj Prev. 1998 Jun;4(2):89–91.

10. Fuller R. Towards a general theory of driver behaviour. Accid Anal Prev. 2005 May;37(3):461–72.

11. Michon JA. A Critical View of Driver Behavior Models: What Do We Know, What Should We Do? In: Evans L, Schwing RC, editors. Human Behavior and Traffic Safety [Internet]. Boston, MA: Springer US; 1985 [cited 2025 Oct 18]. p. 485–524. Available from: http://link.springer.com/10.1007/978-1-4613-2173-6_19

12. Savoie C, Voyer P, Lavallière M, Bouchard S. Transition from driving to driving-cessation: experience of older persons and caregivers: a descriptive qualitative design. BMC Geriatr. 2024 Mar 4;24(1):219.

13. Nemes G. Új, közlekedésbiztonsági fejlesztéseket sürgető jelenség: növekvő számú idős vezető a közutakon. Közlekedéstudományi Szle. 2013;LXIII. évfolyam(5. szám):9–19.

14. Chihuri S, Mielenz TJ, DiMaggio CJ, Betz ME, DiGuiseppi C, Jones VC, Li G. Driving Cessation and Health Outcomes in Older Adults. J Am Geriatr Soc. 2016 Feb;64(2):332–41.

15. Karthaus M, Falkenstein M. Functional Changes and Driving Performance in Older Drivers: Assessment and Interventions. Geriatrics. 2016 May 20;1(2):12.

16. Depestele S, Ross V, Verstraelen S, Brijs K, Brijs T, Van Dun K, Meesen R. The impact of cognitive functioning on driving performance of older persons in comparison to younger age groups: A systematic review. Transp Res Part F Traffic Psychol Behav. 2020 Aug;73:433–52.

17. Sluimer JD, Van Der Flier WM, Karas GB, Fox NC, Scheltens P, Barkhof F, Vrenken H. Whole-Brain Atrophy Rate and Cognitive Decline: Longitudinal MR Study of Memory Clinic Patients. Radiology. 2008 Aug;248(2):590–8.

18. Salat DH, Kaye JA, Janowsky JS. Prefrontal Gray and White Matter Volumes in Healthy Aging and Alzheimer Disease. Arch Neurol. 1999 Mar 1;56(3):338.

19. Gunning-Dixon FM, Brickman AM, Cheng JC, Alexopoulos GS. Aging of cerebral white matter: a review of MRI findings. Int J Geriatr Psychiatry. 2009 Feb;24(2):109–17.

20. Siren A, Haustein S. Driving licences and medical screening in old age: Review of literature and European licensing policies. J Transp Health. 2015 Mar;2(1):68–78.

21. Dobbs AR, Heller RB, Schopflocher D. A comparative approach to identify unsafe older drivers. Accid Anal Prev. 1998 May;30(3):363–70.

22. Lombardi DA, Horrey WJ, Courtney TK. Age-related differences in fatal intersection crashes in the United States. Accid Anal Prev. 2017 Feb;99:20–9.

23. Wheatley CJ, Carr DB, Marottoli RA. Consensus Statements on Driving for Persons with Dementia. Occup Ther Health Care. 2014 Apr 3;28(2):132–9.

24. Stokin GB, Krell-Roesch J, Petersen RC, Geda YE. Mild Neurocognitive Disorder: An Old Wine in a New Bottle. Harv Rev Psychiatry. 2015 Sep;23(5):368–76.

25. Petersen RC, Lopez O, Armstrong MJ, Getchius TSD, Ganguli M, Gloss D, Gronseth GS, Marson D, Pringsheim T, Day GS, Sager M, Stevens J, Rae-Grant A. Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology. 2018 Jan 16;90(3):126–35.

26. Eramudugolla R, Huque MH, Wood J, Anstey KJ. On-Road Behavior in Older Drivers With Mild Cognitive Impairment. J Am Med Dir Assoc. 2021 Feb;22(2):399-405.e1.

27. Chee JN, Rapoport MJ, Molnar F, Herrmann N, O'Neill D, Marottoli R, Mitchell S, Tant M, Dow J, Ayotte D, Lanctôt KL, McFadden R, Taylor JP, Donaghy PC, Olsen K, Classen S, Elzohairy Y, Carr DB. Update on the Risk of Motor Vehicle Collision or Driving Impairment with Dementia: A Collaborative International Systematic Review and Meta-Analysis. Am J Geriatr Psychiatry. 2017 Dec;25(12):1376–90.

28. Foley DJ, Masaki KH, Ross GW, Md LRW. Driving Cessation in Older Men with Incident Dementia. J Am Geriatr Soc. 2000 Aug;48(8):928–30.

29. Tondelli M, Galli C, Vinceti G, Fiondella L, Salemme S, Carbone C, Molinari MA, Chiari A, Zamboni G. Anosognosia in Early- and Late-Onset Dementia and Its

Association With Neuropsychiatric Symptoms. Front Psychiatry. 2021 May 13;12:658934.

30. Feigin VL, Brainin M, Norrving B, Martins SO, Pandian J, Lindsay P, Grupper FM, Rautalin I. World Stroke Organization: Global Stroke Fact Sheet 2025. Int J Stroke. 2025 Feb;20(2):132–44.

31. Almosallam A, Qureshi AZ, Ullah S, Alibrahim A. Return to driving post stroke; patients' perspectives and challenges in Saudi Arabia. Top Stroke Rehabil. 2022 Apr 3;29(3):192–200.

32. Bellagamba D, Vionnet L, Margot-Cattin I, Vaucher P. Standardized on-road tests assessing fitness-to-drive in people with cognitive impairments: A systematic review. Useche SA, editor. PLOS ONE. 2020 May 18;15(5):e0233125.

33. Ranchet M, Akinwuntan AE, Tant M, Salch A, Neal E, Devos H. Fitness-to-drive agreements after stroke: medical versus practical recommendations. Eur J Neurol. 2016 Sep;23(9):1408–14.

34. Sexton E, McLoughlin A, Williams DJ, Merriman NA, Donnelly N, Rohde D, Hickey A, Wren MA, Bennett K. Systematic review and meta-analysis of the prevalence of cognitive impairment no dementia in the first year post-stroke. Eur Stroke J. 2019 Jun;4(2):160–71.

35. Groeger JA, Murphy G. Driving and cognitive function in people with stroke and healthy age-matched controls. Neuropsychol Rehabil. 2022 Jul 3;32(6):1075–98.

36. Devos H, Akinwuntan AE, Nieuwboer A, Truijen S, Tant M, De Weerdt W. Screening for fitness to drive after stroke: A systematic review and meta-analysis. Neurology. 2011 Feb 22;76(8):747–56.

37. Dziezuk E, Le Boterff Q, Van Ravestyn C, Mairesse S, Hamdoun S, Calvet D, Daghsen L, Carment L, Rosso C, Maier MA, Turc G, Mas JL, Lindberg GP. Divided attention and manual visuomotor control in stroke: a combined dual-task and eye movement study. J NeuroEngineering Rehabil. 2025 Oct 10;22(1):212.

38. Asimakopulos J, Boychuck Z, Sondergaard D, Poulin V, Ménard I, Korner-Bitensky N. Assessing executive function in relation to fitness to drive: A review of tools and their ability to predict safe driving. Aust Occup Ther J. 2012 Dec;59(6):402–27.

39. Hird MA, Vetivelu A, Saposnik G, Schweizer TA. Cognitive, On-road, and Simulator-based Driving Assessment after Stroke. J Stroke Cerebrovasc Dis. 2014 Nov;23(10):2654–70.

40. Starkey NJ, Isler RB. The role of executive function, personality and attitudes to risks in explaining self-reported driving behaviour in adolescent and adult male drivers. Transp Res Part F Traffic Psychol Behav. 2016 Apr;38:127–36.

41. Marshall SC, Molnar F, Man-Son-Hing M, Blair R, Brosseau L, Finestone HM, Lamothe C, Korner-Bitensky N, Wilson KG. Predictors of Driving Ability Following Stroke: A Systematic Review. Top Stroke Rehabil. 2007 Jan;14(1):98–114.

42. Blane A, Falkmer T, Lee HC, Dukic Willstrand T. Investigating cognitive ability and self-reported driving performance of post-stroke adults in a driving simulator. Top Stroke Rehabil. 2018 Jan 2;25(1):44–53.

43. Gasne C, Ranchet M, Evennou M, Delebecque MC, Hoang I, Coignard P, Paire-Ficout L. Fitness-to-drive recommendations in post-stroke patients: a retrospective study. J Stroke Cerebrovasc Dis. 2024 May;107781.

44. Williams LJ, Loetscher T, Hillier S, Hreha K, Jones J, Bowen A, Kernot J. Identifying spatial neglect – an updated systematic review of the psychometric properties of assessment tools in adults post-stroke. Neuropsychol Rehabil. 2025 Mar 16;35(3):628–67.

45. Tavaszi I, Szabó G, Erdősi P, Shenker B, Fazekas G. Application of computerised interactive devices for stroke patients with hemispatial neglect. Ideggyógy Szle. 2025;78(3–4):107–21.

46. Raman A, Faria AV, Colavito M, Hillis AE. Novel Insights into Viewer-Centered Versus Stimulus-Centered Hemispatial Neglect: A Cross-Sectional Behavioral and Imaging Study of Acute Stroke. Brain Sci. 2025 Feb 17;15(2):208.

47. Caggiano P, Jehkonen M. The 'Neglected' Personal Neglect. Neuropsychol Rev. 2018 Dec;28(4):417–35.

48. Bartolomeo P. Visual and motor neglect: Clinical and neurocognitive aspects. Rev Neurol (Paris). 2021 Jun;177(6):619–26.

49. Fisher G, Quel De Oliveira C, Verhagen A, Gandevia S, Kennedy D. Proprioceptive impairment in unilateral neglect after stroke: A systematic review. SAGE Open Med. 2020 Jan;8:2050312120951073.

50. Tavaszi I, Nagy AS, Szabo G, Fazekas G. Neglect syndrome in post-stroke conditions: assessment and treatment (scoping review). Int J Rehabil Res. 2021 Mar;44(1):3–14.

51. Bonato M, Deouell LY. Hemispatial Neglect: Computer-Based Testing Allows More Sensitive Quantification of Attentional Disorders and Recovery and Might Lead to Better Evaluation of Rehabilitation. Front Hum Neurosci [Internet]. 2013 [cited 2025 Jul 30];7. Available from: http://journal.frontiersin.org/article/10.3389/fnhum.2013.00162/abstract

52. Terruzzi S, Albini F, Massetti G, Etzi R, Gallace A, Vallar G. The Neuropsychological Assessment of Unilateral Spatial Neglect Through Computerized and Virtual Reality Tools: A Scoping Review. Neuropsychol Rev. 2024 Jun;34(2):363–401.

53. Deouell LY, Sacher Y, Soroker N. Assessment of spatial attention after brain damage with a dynamic reaction time test. J Int Neuropsychol Soc [Internet]. 2005 Oct [cited 2024 Jun 6];11(06). Available from: http://www.journals.cambridge.org/abstract_S1355617705050824

54. Hillis AE. Developments in treating the nonmotor symptoms of stroke. Expert Rev Neurother. 2020 Jun 2;20(6):567–76.

55. Schulz P, Schaebitz W, Driessen M, Beblo T, Toepper M. Recovery of driving fitness after stroke: A matter of time? J Am Geriatr Soc. 2022 May;70(5):1578–80.

56. Devos H, Hawley CA, Conn AM, Marshall SC, Akinwuntan AE. Driving After Stroke. In: Platz T, editor. Clinical Pathways in Stroke Rehabilitation [Internet]. Cham:

Springer International Publishing; 2021 [cited 2024 May 15]. p. 243–60. Available from: https://link.springer.com/10.1007/978-3-030-58505-1_13

57. Johnston BJ, Mackenzie L, O'Donnell JM, Wesson J. Cognitive clinical assessments to predict performance on an on-road assessment: a scoping review. Disabil Rehabil. 2025 Jun 3;1–9.

58. Molnar FJ, Patel A, Marshall SC, Man-Son-Hing M, Wilson KG. Clinical Utility of Office-Based Cognitive Predictors of Fitness to Drive in Persons with Dementia: A Systematic Review. J Am Geriatr Soc. 2006 Dec;54(12):1809–24.

59. Andriola C, Di Rado GR, Presta García DS, Nodari CT. Classification of driving simulators validation: A case study using an immersive driving simulator. Accid Anal Prev. 2025 Apr;213:107944.

60. Cao S, Samuel S, Murzello Y, Ding W, Zhang X, Niu J. Hazard Perception in Driving: A Systematic Literature Review. Transp Res Rec J Transp Res Board. 2022 Dec;2676(12):666–90.

61. Sawada T, Tomori K, Hamana H, Ohno K, Seike Y, Igari Y, Fujita Y. Reliability and validity of on-road driving tests in vulnerable adults: a systematic review. Int J Rehabil Res. 2019 Dec;42(4):289–99.

62. Sawada T, Sakaue K, Kondou J, Higashikawa Y, Ohno K, Tomori K. Discriminant Validity of the Standardized On-Road Assessment for Drivers (SOAD) Among Stroke Patients in Japan. Cureus [Internet]. 2024 Dec 5 [cited 2025 Oct 19]; Available from: https://www.cureus.com/articles/322584-discriminant-validity-of-the-standardized-on-road-assessment-for-drivers-soad-among-stroke-patients-in-japan

63. Yamin S, Ranger V, Stinchcombe A, Knoefel F, Gagnon S, Bédard M. Using Serial Trichotomization with Neuropsychological Measures to Inform Clinical Decisions on Fitness-to-Drive among Older Adults with Cognitive Impairment. Occup Ther Health Care. 2024 Jan 2;38(1):5–25.

64. Szabó G, Biró I, Udvardi V, Fazekas G. Az idős korú gépjárművezetők kognitív alkalmasságának megítélése: kihívások és megoldási lehetőségek. Ideggyógy Szle. 2025;78(7–8):229–38.

65. Bédard M, Maxwell H, Dubois S, Schurr S, Swoluk C, Colosimo A, Cummings S, Weaver B, Stinchcombe A. Serial Trichotomization to Determine Fitness to Drive: Results From a Cohort of Clients Referred to a Neurology Program. Am J Occup Ther. 2025 May 1;79(3):7903205030.

66. Gibbons C, Smith N, Middleton R, Clack J, Weaver B, Dubois S, Bédard M. Using Serial Trichotomization With Common Cognitive Tests to Screen for Fitness to Drive. Am J Occup Ther. 2017 Mar 1;71(2):7102260010p1–8.

67. Krasniuk S, Crizzle AM. Using Serial Trichotomization to Determine Fitness to Drive in Medically At-Risk Drivers. Am J Occup Ther. 2024 Jan 1;78(1):7801205020.

68. Petzschner FH. Practical challenges for precision medicine. Science. 2024 Jan 12;383(6679):149–50.

69. Nouri F, Lincoln N. Validation of a cognitive assessment: predicting driving performance after stroke. Clin Rehabil. 1992 Nov;6(4):275–81.

70. Selander H, Johansson K, Lundberg C, Falkmer T. The Nordic Stroke Driver Screening Assessment as predictor for the outcome of an on-road test. Scand J Occup Ther. 2010 Jan;17(1):10–7.

71. Yang Y, Khorshidi HA, Aickelin U. A review on over-sampling techniques in classification of multi-class imbalanced datasets: insights for medical problems. Front Digit Health. 2024 Jul 26;6:1430245.

72. Hoggarth PA, Innes CRH, Dalrymple-Alford JC, Jones RD. Prediction of driving ability: Are we building valid models? Accid Anal Prev. 2015 Apr;77:29–34.

73. Branch S, Espinosa N, McKinnon AC, Bennett JM. The relationship between brain structure and function and driving performance in older adults with and without

cognitive impairment: A systematic review. Transp Res Part F Traffic Psychol Behav. 2025 Feb;109:1523–41.

74. Wallis KA, Matthews J, Spurling GK. Assessing fitness to drive in older people: the need for an evidence-based toolkit in general practice. Med J Aust. 2020 May;212(9):396.

75. Szabó G, Pintér J, Molontay R, Fazekas G. Driving after stroke: A trichotomous logistic regression model to support decision making in uncertain cases. J Stroke Cerebrovasc Dis. 2025 Nov;34(11):108439.

76. Molenberghs P, Sale MV. Testing for Spatial Neglect with Line Bisection and Target Cancellation: Are Both Tasks Really Unrelated? Sirigu A, editor. PLoS ONE. 2011 Jul 28;6(7):e23017.

77. Mancuso M, Damora A, Abbruzzese L, Navarrete E, Basagni B, Galardi G, Caputo M, Bartalini B, Bartolo M, Zucchella C, Carboncini MC, Dei S, Zoccolotti P, Antonucci G, De Tanti A. A New Standardization of the Bells Test: An Italian Multi-Center Normative Study. Front Psychol. 2019 Jan 22;9:2745.

78. Tsiakiri A, Christidi F, Tsiptsios D, Vlotinou P, Kitmeridou S, Bebeletsi P, Kokkotis C, Serdari A, Tsamakis K, Aggelousis N, Vadikolias K. Processing Speed and Attentional Shift/Mental Flexibility in Patients with Stroke: A Comprehensive Review on the Trail Making Test in Stroke Studies. Neurol Int. 2024 Jan 23;16(1):210–25.

79. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. Nat Protoc. 2006 Dec;1(5):2277–81.

80. Krishnasamy C, Unsworth CA. Normative data, preliminary inter-rater reliability and predictive validity of the Drive Home Maze Test. Clin Rehabil. 2011 Jan;25(1):88–95.

81. Unsworth CA, Baker A, Lannin N, Harries P, Strahan J, Browne M. Predicting fitness-to-drive following stroke using the Occupational Therapy – Driver Off Road Assessment Battery. Disabil Rehabil. 2019 Jul 17;41(15):1797–802.

82. Scarpina F, Tagini S. The Stroop Color and Word Test. Front Psychol [Internet]. 2017 Apr 12 [cited 2025 May 6];8. Available from: http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00557/full

83. Müller VI, Cieslik EC, Ficco L, Tyralla S, Sepehry AA, Aziz-Safaie T, Feng C, Eickhoff SB, Langner R. Not All Stroop-Type Tasks Are Alike: Assessing the Impact of Stimulus Material, Task Design, and Cognitive Demand via Meta-analyses Across Neuroimaging Studies. Neuropsychol Rev [Internet]. 2024 Sep 12 [cited 2025 Oct 19]; Available from: https://link.springer.com/10.1007/s11065-024-09647-1

84. Unsworth CA, Pallant JF, Russell KJ, Germano C, Odell M. Validation of a Test of Road Law and Road Craft Knowledge With Older or Functionally Impaired Drivers. Am J Occup Ther. 2010 Mar 1;64(2):306–15.

85. Statsenko Y, Habuza T, Gorkom KNV, Zaki N, Almansoori TM. Applying the Inverse Efficiency Score to Visual–Motor Task for Studying Speed-Accuracy Performance While Aging. Front Aging Neurosci. 2020 Dec 9;12:574401.

86. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression [Internet]. 1st ed. Wiley; 2013 [cited 2025 May 5]. (Wiley Series in Probability and Statistics). Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118548387

87. Rühl J, Friedrich S. Resampling-based confidence intervals and bands for the average treatment effect in observational studies with competing risks. Stat Comput. 2024 Jun;34(3):101.

88. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clin Transl Sci. 2024 Nov;17(11):e70056.

89. Tefft BC. Rates of Motor Vehicle Crashes, Injuries and Deaths in Relation to Driver Age, United States, 2014-2015 (Research Brief). Wash DC AAA Found Traffic Saf.

90. Hole G. The Effects of Age on Driving. In: The Psychology of Driving [Internet]. 1st ed. Routledge; 2018 [cited 2024 Apr 12]. p. 45–58. Available from:

https://www.taylorfrancis.com/books/9781315516523/chapters/10.4324/9781315516
530-4

91. Ichikawa M, Inada H, Nakahara S. Revisiting Older Drivers' Risks of At-fault Motor Vehicle Collisions in Japan. J Epidemiol. 2023;JE20230217.

92. Siren A, Meng A. Cognitive screening of older drivers does not produce safety benefits. Accid Anal Prev. 2012 Mar;45:634–8.

93. Hakamies-Blomqvist L, Johansson K, Lundberg C. Medical Screening of Older Drivers as a Traffic Safety Measure—A Comparative Finnish-Swedish Evaluation Study. J Am Geriatr Soc. 1996 Jun;44(6):650–3.

94. Langford J, Fitzharris M, Newstead S, Koppel S. Some consequences of different older driver licensing procedures in Australia. Accid Anal Prev. 2004 Nov;36(6):993–1001.

95. Kim JH. Multicollinearity and misleading statistical results. Korean J Anesthesiol. 2019 Dec 1;72(6):558–69.

96. Henézi D, Horváth B, Szegedi A. Idősek közlekedésbiztonsága. Közlekedéstudományi Szle. 2017;67(2):72–8.

97. Berndt A, Clark M, May E. Dementia severity and on-road assessment: Briefly revisited. Australas J Ageing. 2008 Sep;27(3):157–60.

98. Vrkljan BH, McGrath CE, Letts LJ. Assessment Tools for Evaluating Fitness to Drive: A Critical Appraisal of Evidence. Can J Occup Ther. 2011 Apr;78(2):80–96.

99. Stefanidis KB, Summers MJ, Clark M. The assessment of cognitive function and medical fitness to drive in older adults: A qualitative study with general practitioners. J Transp Health. 2025 Jun;42:102020.

100. Bédard M, Maxwell H, Gibbons C, Dubois S, Weaver B, Middleton R. A Three-Tiered Comprehensive Driving Evaluation Integrating a Driving Simulator Test for

Drivers with Borderline Cognitive Fitness-to-Drive: Proof of Concept. Occup Ther Health Care. 2025 Jul 3;39(3):479–99.

101. Piersma D, Fuermaier ABM, De Waard D, Davidse RJ, De Groot J, Doumen MJA, Bredewoud RA, Claesen R, Lemstra AW, Scheltens P, Vermeeren A, Ponds R, Verhey F, De Deyn PP, Brouwer WH, Tucha O. Assessing Fitness to Drive in Patients With Different Types of Dementia. Alzheimer Dis Assoc Disord. 2018 Jan;32(1):70–5.

102. Kay LG, Bundy AC, Clemson LM. Predicting Fitness to Drive in People With Cognitive Impairments by Using DriveSafe and DriveAware. Arch Phys Med Rehabil. 2009 Sep;90(9):1514–22.

103. Dobbs BM, Schopflocher D. The Introduction of a New Screening Tool for the Identification of Cognitively Impaired Medically At-Risk Drivers: The SIMARD A Modification of the DemTect. J Prim Care Community Health. 2010 Jul;1(2):119–27.

104. Nakas CT, Bantis LE, Gatsonis CA. ROC Analysis for Classification and Prediction in Practice [Internet]. 1st ed. Boca Raton: Chapman and Hall/CRC; 2023 [cited 2025 Oct 19]. Available from: https://www.taylorfrancis.com/books/9780429170140

105. Geroldinger A, Lusa L, Nold M, Heinze G. Leave-one-out cross-validation, penalization, and differential bias of some prediction model performance measures—a simulation study. Diagn Progn Res. 2023 May 2;7(1):9.

106. Samuelsson K, Wressle E. Usability of Norm Values on Results from a Simulator Device and Cognitive Tests in Traffic Medicine. Med Devices Evid Res. 2022 Mar;Volume 15:59–70.

107. Quintas JL, Trindade IOA, Gameiro KSD, Pitta LSR, Camargos EF, Nóbrega OT. Neuropsychological domains and fitness to drive in mild cognitive impairment or Alzheimer's disease. Accid Anal Prev. 2023 Oct;191:107188.

108. Rike PO, Johansen HJ, Ulleberg P, Lundqvist A, Schanke AK. Exploring associations between self-regulatory mechanisms and neuropsychological functioning and driver behaviour after brain injury. Neuropsychol Rehabil. 2018 Apr 3;28(3):466–90.

109. Ghawami H, Homaei Shoaa J, Moazenzadeh M, Sorkhavandi M, Okhovvat A, Hadizadeh N, Yamola M, Rahimi-Movaghar V. Ecological validity of executive function tests in predicting driving performance. Appl Neuropsychol Adult. 2024 Nov;31(6):1352–64.

110. Li S, Qian D, Li P, Yuan X, Fang Q. Risk prediction model for distracted driving: Characterizing interactions of eye glances and manual sequences. Travel Behav Soc. 2024 Oct;37:100851.

111. Uc EY, Rizzo M, Anderson SW, Qian S, Rodnitzky RL, Dawson JD. Visual dysfunction in Parkinson disease without dementia. Neurology. 2005 Dec 27;65(12):1907–13.

112. Dawson JD, Anderson SW, Uc EY, Dastrup E, Rizzo M. Predictors of driving safety in early Alzheimer disease. Neurology. 2009 Feb 10;72(6):521–7.

113. Lundberg C, Caneman G, Samuelsson S, Hakamies-Blomqvist L, Almkvist O. The assessment of fitness to drive after a stroke: The Nordic Stroke Driver Screening Assessment. Scand J Psychol. 2003 Feb;44(1):23–30.

114. Tan KM, O'Driscoll A, O'Neill D. Factors affecting return to driving post-stroke. Ir J Med Sci. 2011 Mar;180(1):41–5.

115. Tant MLM, Brouwer WH, Cornelissen FW, Kooijman AC. Driving and visuospatial performance in people with hemianopia. Neuropsychol Rehabil. 2002 Nov;12(5):419–37.

116. Kooijman A, Brouwer W, Coeckelbergh T, Tant M, Cornelissen F, Bredewoud R, Melis-Dankers B. Compensatory viewing training improves practical fitness to drive of subjects with impaired vision. Vis Impair Res. 2004 Jan;6(1):1–27.

117. Sotokawa T, Nasu S, Ikuta J, Sonohara K. Evaluation of driving fitness using driving simulators in patients with right-hemisphere damage: an unmatched case-control study. Top Stroke Rehabil. 2024 Feb 17;31(2):167–77.

118. Wood JM, McGwin G, Elgin J, Vaphiades MS, Braswell RA, DeCarlo DK, Kline LB, Meek GC, Searcey K, Owsley C. On-Road Driving Performance by Persons with Hemianopia and Quadrantanopia. Investig Opthalmology Vis Sci. 2009 Feb 1;50(2):577.

119. Bro T, Andersson J. Individuals with severe visual field loss from stroke and glaucoma could have on-road driving safety comparable to normally sighted drivers. Acta Ophthalmol (Copenh). 2025 Nov;103(7):842–9.

120. Overman MJ, Binns E, Milosevich ET, Demeyere N. Recovery of Visuospatial Neglect With Standard Treatment: A Systematic Review and Meta-Analysis. Stroke. 2024 Sep;55(9):2325–39.

121. Azouvi P, Rousseaux M, Bartolomeo P, Pérennou D, Pradat-Diehl P, Wiart L, Rode G, Godefroy O. Discriminative value of different combinations of tests to detect unilateral neglect in patients with right hemisphere damage. Eur J Neurol. 2023 Oct;30(10):3332–40.

122. Andres M, Geers L, Marnette S, Coyette F, Bonato M, Priftis K, Masson N. Increased Cognitive Load Reveals Unilateral Neglect and Altitudinal Extinction in Chronic Stroke. J Int Neuropsychol Soc. 2019 Jul;25(6):644–53.

123. Ten Brink AF, Verwer JH, Biesbroek JM, Visser-Meily JMA, Nijboer TCW. Differences between left- and right-sided neglect revisited: A large cohort study across multiple domains. J Clin Exp Neuropsychol. 2017 Aug 9;39(7):707–23.

124. Gottesman RF, Kleinman JT, Davis C, Heidler-Gary J, Newhart M, Kannan V, Hillis AE. Unilateral neglect is more severe and common in older patients with right hemispheric stroke. Neurology. 2008 Oct 28;71(18):1439–44.

125. Walle KM, Nordvik JE, Becker F, Espeseth T, Sneve MH, Laeng B. Unilateral neglect post stroke: Eye movement frequencies indicate directional hypokinesia while fixation distributions suggest compensational mechanism. Brain Behav. 2019 Jan;9(1):e01170.

126. Takamura Y, Imanishi M, Osaka M, Ohmatsu S, Tominaga T, Yamanaka K, Morioka S, Kawashima N. Intentional gaze shift to neglected space: a compensatory strategy during recovery after unilateral spatial neglect. Brain. 2016 Nov;139(11):2970–82.

127. Giannakou I, Lin D, Punt D. Computer-based assessment of unilateral spatial neglect: A systematic review. Front Neurosci. 2022 Aug 19;16:912626.

128. Holden A, Pusey H. The impact of driving cessation for people with dementia – An integrative review. Dementia. 2021 Apr;20(3):1105–23.

9. Bibliography of the candidate's publications

9.1. Publications related to the dissertation

- Szabó G, Pintér J, Molontay R, Fazekas G. Driving after stroke: A trichotomous logistic regression model to support decision making in uncertain cases. Journal of Stroke and Cerebrovascular Diseases. 2025;34(11):108439. **IF: 1.8**

- Szabó G, Biró I, Udvardi V, Fazekas G. Az idős korú gépjárművezetők kognitív alkalmasságának megítélése: kihívások és megoldási lehetőségek. Ideggyógyászati Szemle. 2025;78(7–8):229–38. **IF: 0.6**

9.2. Publications not directly related to the dissertation

- Tavaszi I, Szabó G, Erdősi P, Shenker B, Fazekas G. Application of computerised interactive devices for stroke patients with hemispatial neglect. Ideggyógyászati szemle. 2025;78(3–4):107–21. **IF: 0.6**

- Udvardi V, Szabo G, Takacs J, Fazekas G. The effectiveness of mindfulness-based cognitive therapy during poststroke rehabilitation: a randomized controlled trial. International Journal of Rehabilitation Research. 2024. szeptember;47(3):169–75. **IF: 1.3**

- Udvardi V, Szabó G, Fazekas G. A tudatos jelenlét alapú beavatkozások hatásai a stroke utáni rehabilitációban. Ideggyógyászati szemle. 2023;76(1–2):11–7. **IF: 0.9**

- Tavaszi I, Nagy AS, Szabo G, Fazekas G. Neglect syndrome in post-stroke conditions: assessment and treatment (scoping review). International Journal of Rehabilitation Research. 2021;44(1):3–14. **IF: 1.832**

∑**IF: 7.032**

10. Acknowledgements