ELSEVIER

Journal of the Neurological Sciences 260 (2007) 188–198

Journal of the Neurological Sciences

www.elsevier.com/locate/jns

Sensory-motor and cognitive tests predict driving ability of persons with brain disorders

Carrie R.H. Innes [a,b,c,*], Richard D. Jones [a,b,c,d], John C. Dalrymple-Alford [a,e], Sarah Hayes [f], Sue Hollobon [f], Julie Severinsen [f], Gwyneth Smith [f], Angela Nicholls [f], Tim J. Anderson [a,b,g]

[a] *Van der Veer Institute for Parkinson's and Brain Research, Christchurch, New Zealand*
[b] *Department of Medicine, Christchurch School of Medicine and Health Sciences, University of Otago, Christchurch, New Zealand*
[c] *Department of Medical Physics and Bioengineering, Christchurch Hospital, Christchurch, New Zealand*
[d] *Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand*
[e] *Department of Psychology, University of Canterbury, Christchurch, New Zealand*
[f] *Department of Occupational Therapy, Burwood Hospital, Christchurch, New Zealand*
[g] *Department of Neurology, Christchurch Hospital, Christchurch, New Zealand*

Received 30 June 2006; received in revised form 23 April 2007; accepted 25 April 2007
Available online 4 June 2007

Abstract

Objective: Brain disorders can lead to a decreased ability to perform the physical and cognitive functions necessary for safe driving. This study aimed to determine how accurately a battery of computerized sensory-motor and cognitive tests (*SMCTests*™) could predict driving abilities in persons with brain disorders.
Methods: SMCTests and an independent on-road driving assessment were applied to 50 experienced drivers with brain disorders referred to a hospital-based driving assessment service. The patients comprised 36 males and 14 females, a mean age of 71.3 years (range 43–85 years) and diagnoses of 35 stroke, 4 traumatic brain injury, 4 Alzheimer's disease, and 7 other. Binary logistic regression (BLR) and nonlinear causal resource analysis (NCRA) were used to build model equations for prediction of on-road driving ability based on *SMCTests* performance.
Results: BLR and NCRA correctly classified 94% and 90% of referrals respectively as on-road pass or fail. Leave-one-out cross-validation estimated that BLR and NCRA would correctly predict the classification of 86% and 76% respectively of an independent referral group as on-road pass or fail.
Conclusions: Compared with other studies, *SMCTests* have shown the highest predictive accuracy against true on-road driving ability as estimated in an independent data set and in persons with brain disorders. *SMCTests* also have the advantage of being able to comprehensively and objectively assess both sensory-motor and higher cognitive functions related to driving.

Keywords: Driving; Brain disorders; Prediction; Sensory-motor; Cognitive; human performance

1. Introduction

Ability to drive is highly valued by most adults and personal transport is used daily for all manner of activities. Unfortunately, a wide variety of brain disorders can lead to impairment of sensory-motor and/or cognitive abilities important for safe driving [1–4]. Thus, it is important that appropriate driving assessment is available to identify persons no longer able to drive safely. For the driver, forfeiture of a driver's licence can have a marked negative effect on the person's mobility, and has been linked to depression, lowered self-esteem, and reduced overall quality of life [5]. Therefore, any driving assessment must be thorough, objective, and strive to ensure that all those who fail do so because of marked deficits that would unquestionably lead to unsafe driving.

* Corresponding author. Van der Veer Institute for Parkinson's and Brain Research, 66 Stewart Street, Christchurch, New Zealand. Tel.: +64 3 378 6096; fax: +64 3 378 6080.
E-mail address: carrie.innes@vanderveer.org.nz (C.R.H. Innes).

0022-510X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jns.2007.04.052

Where possible, the assessment should also provide failed or borderline subjects with specific information on, and potential consequences of, their deficits, in order to obtain as high a level as possible of acceptance of, and compliance with, assessment and rehabilitation recommendations.

In New Zealand, medical practitioners have a legal obligation to consider medical aspects of fitness to drive when conducting a medical examination to determine if an individual is fit to drive [6]. In cases where there is doubt regarding an individual's fitness to drive, a practitioner can refer a patient to a specialist driving assessment service. Occupational therapists with skills in driver assessment offer services to meet this need in major centres in New Zealand. Approximately 1400 persons (35 per 100,000) with brain lesions or age-related cognitive decline are referred for a driving assessment in New Zealand (population ~4 million) each year. Despite a relatively high demand for specialized driving assessment, an accurate, comprehensive, and standardized off-road assessment procedure has yet to be identified.

Internationally, several groups have investigated the relationship between on-road driving and performance on neuropsychological test batteries [7–16], tests specifically designed for driving-related functions [17–23], and driving simulators [9,24–27]. Researchers have also investigated the relationship between off-road test performance and on-road driving as estimated by crash records [28–31], closed-course driving [32,33], or simulated driving [34–38].

Several of these studies have investigated the relationship between off- and on-road assessment outcome (i.e., Pass or Fail) using classification models, in which the test data is the same as the training data. Classification models have proven to be 70–94% accurate in classifying on-road assessment outcome based on performance on off-road tests [7–9,15–17,22,27] or driving simulators [27]. A more robust determination of the relationship between off-road assessments and on-road driving outcome can be achieved by studies using predictive models, in which models are trained and tested on separate data groups. Predictive models have proven to be 62–89% accurate in predicting on-road assessment outcome in an independent group of subjects [14,19,23,32,39].

In addition to varying success at identifying at-risk drivers, there is a need for a simple, relatively inexpensive, but comprehensive assessment system that covers both the physical and cognitive deficits underlying an inability to drive safely in certain individuals with neurological disorders. Such an assessment system would need to have a high predictive accuracy for on-road driving ability. We have developed a battery of computerized sensory-motor and cognitive tests (*SMCTests*™) to quantify sensory-motor and cognitive dysfunction as a research and quantitative assessment tool in neurology and neurorehabilitation, with particular application to the assessment of driving abilities in patients with neurological disorders. The sensory-motor tests comprise tests of visuoperception, ballistic movements, and visuomotor tracking which have already been extensively used as a standalone battery [40,41] to assess sensory-motor dysfunction in off-road driving assessment [42,43], unilateral stroke [44,45], Parkinson's disease [46–49], stutterers [50], and following alcohol consumption [51]. The cognitive tests were newly developed measures of complex attention, visual search, decision-making, impulse control, planning, and divided attention. These cognitive functions were identified following an analysis of the literature describing (a) perceived cognitive deficits observed during on-road assessment of subjects with brain disorders (e.g., [1–4]) and (b) performance on cognitive tests associated with driving in subjects with brain disorders [8–23,28–30,32,35,36]. The tests were designed to be contextually close to the on-road driving task so as to provide higher face validity than readily available standard neuropsychological tests, to increase user compliance, and encourage optimal prediction of on-road performance. The normative range, reliability, validity, and effect of sex and age on performance of the cognitive tests have been determined in a separate study of 60 healthy control subjects (manuscript in preparation). This showed that, compared with standard tests of related cognitive functions, the new cognitive tests achieved good construct reliability, good test–retest reliability, and subjectively higher face validity. There was a trend for age to have a detrimental effect on performance but sex differences were seen on only one test, with males having superior performance on a 'planning' test.

The primary objective of this study was to determine the extent to which performance on *SMCTests* can be used to predict on-road driving ability in subjects with brain disorders. An accurate prediction of on-road driving would in large part minimize the need for on-road driving assessments for referrals who will ultimately fail, thus decreasing the unnecessary cost and risk of accident these entail. Reducing the proportion of referrals needing an on-road assessment will lead to lower costs or a higher throughput of referrals for the same resources and costs. In addition, a comprehensive evaluation of cognitive and sensory-motor functions will help identify physical, perceptual, and cognitive deficits underlying an inability to drive safely and allow focussed rehabilitation where possible.

2. Methods

2.1. Subjects

A consecutive sample of 50 subjects with brain disorders who had been referred to the Driving and Vehicle Assessment Service (DAVAS) at Burwood Hospital, Christchurch, were recruited to the study over a 14 month period. All subjects wished to return to, or continue, driving despite a medical condition that might have affected their driving ability. Subjects had been referred to DAVAS through general practitioners, the New Zealand Accident Compensation Corporation, or District Health Board practitioners. All referrals had a definite or probable brain disorder as follows: 35 stroke, 4 traumatic brain injury, 4 Alzheimer's disease, 2 age-related cognitive decline, 1 multiple sclerosis, 1 Parkinson's disease, 1 hypoxia, and 2 other neurological disorders. They were

required to be free from any unrelated diagnosed psychiatric illness and referrals unable to use their lower limbs to drive were excluded from the study. Referrals either held a current driver's licence or had held one prior to the brain disorder (mean of 50.5 years driving experience, range 12–69 years). Twenty-six referrals were actively driving at the time of assessment, while 24 referrals had not driven for a mean of 6.2 months prior to assessment (range 1–12 months). There were 36 males and 14 females with a mean age of 71.3 years (range 43–85 years). A further 204 referrals were eligible for the study but declined to participate. The 50 referrals that were tested were closely representative of sex, age, and neurological disorders of the overall group.

To provide a comparison group, 12 healthy control subjects were also recruited to the study. This group comprised two females and two males from each of three age-range groups (20–40 years, 41–60 years, and 61–80 years). Controls had a mean of 34 years driving experience (range 12–57 years). None of the controls had suffered any form of brain injury or had any diagnosed psychiatric, neurological, or musculoskeletal disorder.

Ethical approval for the study was obtained through the New Zealand Canterbury Ethics Committee prior to recruitment and testing.

2.2. Apparatus

Subjects were tested in a modified car body interfaced to a Pentium PC (Fig. 1). They used the steering wheel, indicator stick, accelerator, clutch, and brake pedals (or hand controls) to respond to 80×60 cm computer-generated test stimuli displayed by a data projector on to a wall in front with an eye-to-screen distance of 180 cm (giving a visual angle of +/−11.3°). The *SMCTests* program, run by an assessor on a separate monitor, generated the tests, analysed performance, stored biographical and test data in a database, and printed performance summaries.

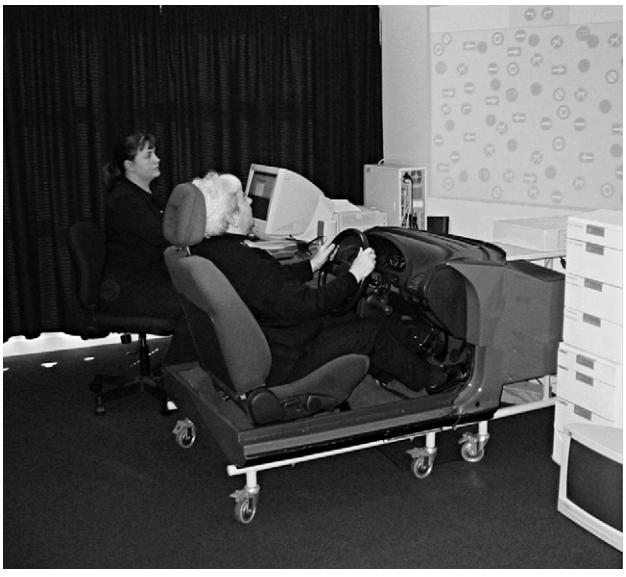

Fig. 1. *SMCTests* apparatus.

2.3. Sensory-Motor tests

The sensory-motor tests are described in detail elsewhere [40,41,44,46,52]. Briefly, they include three visuoperceptual tests (Visual Resolution, Static Perception, and Dynamic Perception), four visuomotor tests (Ballistic Movement, Footbrake Reaction, Footbrake and Clutch Reaction, and Hand Control Reaction), and three eye-arm tracking tests (Sine Tracking, Random Tracking and Step Tracking) which involve the use of a steering wheel and/or foot pedals. Visual Resolution measured the minimum separation at which a subject is able to identify a dot as being off the centre of a vertical line. Static Perception measured the minimum separation at which a subject is able to identify the tip of an arrow as being off a vertical line and a sinusoidal waveform. Dynamic Perception measured the minimum separation between the point of an arrow and a moving random preview target at which a subject is able to perceive the tip of an arrow as being off the target (as in Fig. 2A but without the 4 horizontally pointing arrows which were used in a subsequent test). Ballistic Movement measured the reaction time and maximum speed at which a subject can turn the steering wheel to move an arrow out of a box and across a pass-line in response to an unexpected signal. Footbrake Reaction measured the reaction and movement times at which a subject can respond to an unexpected signal by moving his or her foot from the accelerator to the brake. Footbrake and Clutch Reaction measured the reaction and movement times at which a subject can respond to an unexpected signal by releasing the accelerator and pressing both the clutch and brake pedals. Hand Control Reaction measured the reaction and movement times at which a subject can respond to an unexpected signal by pushing back on a hand control lever. Tracking measured the accuracy with which a subject can track a laterally moving target using the steering wheel to move an on-screen arrow. The tracking target may move in a sine wave (Sine Tracking), random wave (Random Tracking) (as in Fig. 2A but without the 4 horizontally pointing arrows), or jump to the left or right (Step Tracking).

2.4. Cognitive tests

2.4.1. Divided Attention (Fig. 2A)

This test evaluated ability to divide attention between two simultaneously performed visuo-cognitive activities. The test combined a preview random tracking task with a simultaneous visual scanning task (Arrows Perception). While the subject tracked the random target (8.0 s preview) with the steering wheel, 12 consecutive sets of four arrows were displayed on the same screen. The subject aimed to maintain accurate tracking of the target while scanning the arrows and determining whether or not all 4 arrows were pointing in the same direction. Each set of arrows was displayed on screen

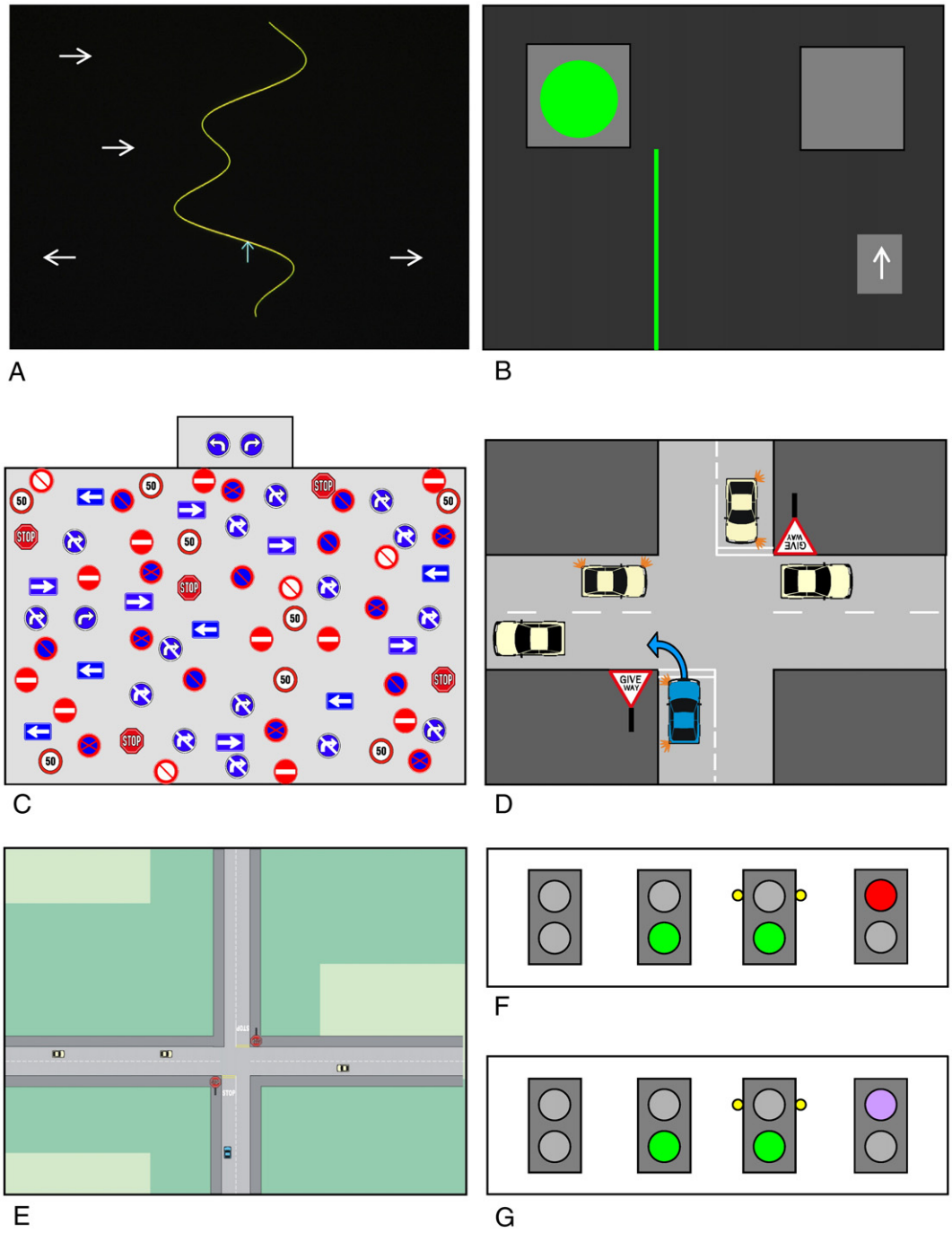

Fig. 2. Sample screens from (A) Divided Attention — The subject must determine whether all four horizontal arrows are pointing in the same direction or not while maintaining accurate tracking of the curve with the tip of the vertical arrow, (B) Complex Attention — Using the steering wheel, the subject must move the arrow from the grey box past the green line as quickly as possible when the green light changes from the right box to the left box, (C) Visual Search — The subject must locate within the large box of 70 stimuli an example of one of the target stimuli shown in the top box. In this case the 'turn-right arrow' is presented and the subject should turn the steering wheel towards the right as quickly as possible, (D) Decision-Making — The subject must decide as quickly as possible whether it is necessary to give way to any other car. In this case, it is safe for the blue car to proceed and the subject should press the accelerator, (E) Planning — The subject must drive the blue car down a road and cross intersections while avoiding hazards and other vehicles, (F) Impulse Control — Four consecutive screens of a standard stimulus trial in which the subject aims to release the accelerator and press the brake as quickly as possible when the red light is presented, (G) Impulse Control — Four consecutive screens of a false stimulus trial in which the subject must suppress any impulse to release the accelerator when the purple light is presented.

for 4.8 s, with a 1.0 s delay between sets. The subject was tested separately on the tracking and arrow perception tasks in order to obtain baseline performance data.

2.4.2. Complex Attention (Fig. 2B)

This test assessed ability to sustain complex attention. The subject used the steering wheel to maintain an arrow in a

box on the same side of the screen as a green light symbol. This light symbol alternated between the left and right side of the screen, requiring the subject to turn the steering wheel from left to right repeatedly. Each green light was presented for 3.0–5.0 s. Reaction, movement, and total times were measured to give indications of slowed information processing, mental and physical fatigue, and lapses in concentration.

2.4.3. Visual Search (Fig. 2C)

This test assessed visual scanning and selective attention, including left-right or central-peripheral vision bias, and comprised 20 trials of static images, each containing 70 road-sign stimuli. Each screen was presented for a maximum of 10 s. The subject searched each image for one of two target stimuli — a 'turn-left arrow' or a 'turn-right arrow'. Only one of the two target stimuli was present on any trial. If the 'turn-left arrow' was presented, the subject was required to turn the steering wheel towards the left as quickly as possible, and conversely for the 'turn-right arrow'. Examples of the target stimuli were shown in a box at the top of all images.

2.4.4. Decision-Making (Fig. 2D)

This test assessed accuracy and speed of decision-making related to road rules. The subject was presented with images of a bird's eye perspective of an intersection involving two or more cars. A blue car near the bottom of the image represented the subject's car, with all other cars being yellow. Orange indicator lights signalled the driving intention of each car. A blue arrow extended off the front of the subject's car to indicate intended direction of travel. For each screen, the subject needed to decide as quickly as possible whether it was necessary to give way to another car and press brake, or whether it was safe to proceed and press accelerator. The 26 trials were split into 12 basic trials (subject's car and one other car) and 14 complex trials (subject's car and 3–6 other cars).

2.4.5. Planning (Fig. 2E)

This test assessed ability to use accurate timing and judgement as an indicator of planning ability. The subject was presented with a screen showing a bird's eye view of a road and surrounds, and told that they were the blue car near the bottom of the screen. The road and surrounds were appropriate for driving on the left-side of the road. When the subject pressed the accelerator, the road environment scrolled down the screen, which simulated the subject's car driving forward along the road. A preview time of 18.5 s represented scaled equivalents of a 257 m preview distance and speed of 50 km/hr. The subject used the brake in order to 'stop' the blue car. The road was primarily straight but included four curved sections. At predetermined intervals a hazard or crossroad appeared in the road ahead. The aim was for subjects to drive as far as possible in 6 min while avoiding all hazards. The subject needed to avoid overtaking hazards or crossing intersections when on-coming traffic blocked the subject's travel.

2.4.6. Impulse Control (Fig. 2F and G)

This test assessed a subject's ability to exercise appropriate anticipatory and inhibitory control. The subject depressed an accelerator pedal to activate a green light (simulating the bottom light in a set of traffic lights) on a screen. When the green light was extinguished, a red light was presented immediately above and the subject needed to release the accelerator pedal and depress the brake pedal as quickly as possible. This was a standard-stimulus trial. In a minority of trials (10 out of 45), a purple rather than a red light was presented. This was a false stimulus trial for which the subject was asked to keep the accelerator pedal depressed. The presentation of the top light in both the standard and false stimulus trials was cued by a separate signal – a yellow circle on either side of the top light – 500 ms prior to the light change. Incorrect releases in response to false stimulus trials were interpreted as the subject exhibiting a deficit in inhibitory control. Incorrect releases of the accelerator pedal while the bottom green light was still presented (or within 180 ms of the light change) were interpreted as the subject exhibiting a deficit in anticipatory control.

2.5. Procedure

The 50 DAVAS referrals and 12 control subjects were assessed off-road on *SMCTests* and, in a subsequent session, on-road. The on-road assessment was conducted by a driving occupational therapist and a driving instructor blinded to performance on *SMCTests*. Both driving instructor and occupational therapist were experienced in driver assessment of persons with disabilities or brain disorders. The driving instructor was seated in the front passenger seat and was responsible for giving directions to the subject and for maintaining the safety of the vehicle, as well as providing an evaluation of performance. All study participants performed the same standardized on-road assessment as DAVAS referrals who declined to participate in the study. The on-road assessment began within the hospital grounds where basic control of the vehicle was assessed, and continued on an open course in a nearby suburb which experienced relatively little traffic but included controlled and uncontrolled intersections. The assessment then continued in increasingly busy and complicated traffic situations. Traffic hazards included dual-lane roads, single-lane roundabouts, dual-lane roundabouts, controlled intersections, uncontrolled intersections, and changes in speed zone. Assessments were approximately 45 min unless the safety of the vehicle, occupants, or other road users was considered at risk at any stage during the assessment. On-road driving performance was independently scored by the experienced occupational therapist and driving instructor as Pass or Fail using the Advanced Driving Assessment System, which is the standard system used by all driving assessment occupational therapists in New Zealand. Advanced Driving Assessment System courses are run through the New Zealand Institute of Driving Instructors. Performance was scored in four main areas: search, hazard

Table 1
Driving Scale

Driving score	Outcome	Label	Detail
0	Fail	No ability	Complete inability to perform the physical and cognitive requirements of driving
1	Fail	Basic skills only	Able to perform only the most basic physical and cognitive requirements of driving
2	Fail	Extremely inferior	Several major and minor errors or difficulties
3	Fail	Very poor	Some major and minor errors or difficulties
4	Fail	Poor	One major error or difficulty, or several minor errors or difficulties not fully rectified upon instruction
5	Fail	Borderline	Some minor errors or difficulties not fully rectified upon instruction (driving lessons/adaptations may be required)
6	Pass	Fair	More than two minor errors or difficulties (rectified with instruction)
7	Pass	Satisfactory	Two minor errors or areas of difficulty (rectified with instruction)
8	Pass	Good	One minor error or area of difficulty (rectified with instruction)
9	Pass	Very good	No errors or difficulties
10	Pass	Flawless	Flawless driving performance — no errors or difficulties and superior skills

identification, controls, and observation of traffic regulations. In addition, on-road performance was also rated on a Driving Scale (0–10) specifically designed for the study (Table 1).

2.6. Data analysis

Analysis of the referral subject performance on *SMCTests* was undertaken to quantify the sensitivity (i.e., correct prediction of on-road assessment fails) and specificity (i.e., correct prediction of on-road assessment passes) of the tests at group and individual levels. As the majority of the data were not normally distributed (Shapiro–Wilk W test, $p<.05$), and several measures were ordinal, non-parametric techniques were used to analyse the data.

At group level, Mann–Whitney U analysis (STATISTICA 6.0, StatSoft, Inc.) was undertaken to determine significant differences in off-road test performance between the referrals who passed the on-road assessment and those who failed. The Cohen effect size statistic for rank-transformed variables [53] was used to evaluate the magnitude of differences in off-road performance between the pass and fail referral groups.

Two methods were used to determine the predictive value of performance on *SMCTests* for on-road driving ability at the individual level for referrals. The first of these, binary logistic regression (BLR), is a non-parametric version of discriminant analysis for the case where the dependent variable is dichotomous (i.e., pass or fail on an on-road assessment). BLR was used to estimate the probability of an on-road assessment fail based on an exponential function of *SMCTests* variables and weightings. A forward stepwise method was used to select the optimal set of *SMCTests* variables for predicting on-road assessment outcome. Forward stepwise

Table 2
SMCTests measures which showed a significant difference in performance between the on-road assessment pass and fail groups based on Mann–Whitney U analysis

Test measure	Healthy subjects * $n=40$ Median	Pass referrals $n=18$ Median	Fail referrals $n=32$ Median	Mann–Whitney U Pass vs Fail Referrals p-value	Effect size of Pass vs Fail **
Complex Attention — reaction and movement time (ms)	857	1031	1388	0.001	1.53
Tracking — sine tracking mean error (mm)	8.4	12.2	18.5	0.001	1.36
Tracking — step tracking mean error (mm)	14.1	14.3	17.2	0.001	1.29
Ballistic Movement — best reaction and movement time (ms)	499	489	549	0.001	1.28
Planning — number of hazards hit	2	2	4	0.001	1.27
Planning — duration of lateral position errors (s)	6	12	38	0.001	1.27
Impulse Control — mean reaction time (ms)	449	480	616	0.001	1.20
Decision-Making — total number of correct responses	23	22	17	0.001	1.10
Tracking — random tracking mean error (mm)	5.9	9.0	13.5	0.001	1.08
Footbrake — reaction and movement time (ms)	622	574	681	0.002	1.07
Hand Control — reaction and movement time (ms)	361	367	405	0.003	0.95
Divided Attention — dual task random tracking error (mm)	8.3	10.6	17.5	0.005	0.92
Ballistic Movement — peak velocity (mm/s)	1128	1212	1001	0.005	0.91
Planning — safety margin at intersections (mm)	25	27	14	0.005	0.86
Impulse Control — number of commissions	3	3	5	0.009	0.81
Planning — number of crashes	1	1	2	0.023	0.70
Visual Search — correct responses	18	15	11	0.034	0.64
Planning — distance travelled (m)	3.80	3.76	3.54	0.036	0.62

* Performance data from a separate study of healthy subjects comprising 20 male and 20 female experienced drivers aged 41–78 years (mean 60.4 years).
** Effect size calculated using a Cohen effect-size statistic for rank-transformed variables [53].

Table 3
Test measures in binary logistic regression model equation

	Test	Measure
1	Planning	Number of hazards hit
2	Complex Attention	Reaction and movement time
3	Tracking	Sine tracking mean error
4	Ballistic Movement	Best reaction time
5	Divided Attention	Dual task random tracking error

selection begins by adding the most statistically significant variable to the regression equation; variables are then added one at a time until all statistically significant variables have been included [54]. Each time a new variable is added the statistical significance of previously added variables is checked using a chi-square test to ensure they still add significantly to the prediction [54].

The second technique, Nonlinear Causal Resource Analysis (NCRA), is a relatively new approach to performance prediction and is based on the resource economic performance modelling constructs of General Systems Performance Theory and the Elemental Resource Model [55]. With NCRA, the minimum resource level required to achieve a given level of performance on a high-level task is determined for each test function and plotted as a resource demand function (RDF) curve [32,55]. RDF curves were created for key performance measures from each of the *SMCTests* tests. A major benefit of NCRA is that it can then determine the specific test function that maximally limited each subject's performance on the high-level driving task. NCRA-predicted scores for each referral were compared with observed Driving Scale scores (0–10) in order to determine the accuracy of the NCRA model predictions.

Leave-one-out cross-validation was used to estimate the true error rate of the predictive models produced by BLR and NCRA. It is a well established method for estimating how well a predictive model will function on an independent test set when a data set is too small to allow separate training and test data [56]. In leave-one-out cross-validation each case is left out in turn while the remaining data is used to train the model. The model is then tested on the single case which was left out [56]. This is a computationally intensive technique as the equation needs to be trained and tested once for every case in the data set (i.e., 50 times in our study). However, this method has the advantage that the greatest possible amount of data can be used to train each model. As long as the population in the training set is representative of the population that the predictive model will be used with, this method provides a sound estimate of the predictive accuracy that would be achieved with a separate test dataset [56].

In addition to the leave-one-out cross-validation of the referral group, an additional independent measure of the estimated value of the predictive models was gained by having the 12 healthy control subjects serve as independent data to test the models.

3. Results

Of the 50 referrals, 32 (64%) failed the on-road driving assessment and 18 (36%) passed. These proportions were not different for the referrals who declined to participate in the study (69% failed, 31% passed, Fisher's Exact Test two-tailed $p=.50$). All referrals who failed the on-road assessment failed due to near-misses with other vehicles or hazards or due to definite safety concerns during their on-road assessment. Of the 24 referrals who were not active drivers at the time of assessment, 16 failed and 8 passed the on-road assessment. These proportions were not different to the pass and fail rates of those who had been actively driving at the time of the assessment (16 failed and 10 passed, Fisher's Exact Test two-tailed $p=.77$).

Mann–Whitney U analysis showed that there were differences in off-road test performance between the group of referrals who passed the on-road assessment and those who failed on performance measures from each test in *SMCTests* except the three visuoperceptual tests (Visual, Static, and Dynamic) (Table 2). Cohen effect sizes for rank-transformed variables ranged from 0.62–1.53 and were highest for measures from cognitive tests of Complex Attention, Planning, Impulse Control, and Decision-Making, and from sensory-motor tests of Tracking and upper-limb Ballistic Movement.

Forward stepwise BLR analysis produced a model with 5 *SMCTests* measures (Table 3) able to correctly classify 47 of the 50 referrals (i.e., 94% accuracy on training data) as an on-road Pass or Fail. The sensitivity of the model was 97% (31/32 correctly classified as Fail) and the specificity was 89% (16/18 correctly classified as Pass). Based on BLR analysis, 26/32 (81%) of referrals who failed the on-road assessment had a high probability (0.80–1.00) of failing the on-road assessment, 9/18 (50%) referrals who passed the on-road assessment had a low probability (0.00–0.05) of failing the on-road assessment, and the remaining referrals had probabilities which fell between the two extremes (Table 4). The correlation between the BLR estimated probability of failing the on-road assessment and observed on-road Driving Scale scores was $r=-0.77$ (Fig. 3).

Leave-one-out cross-validation analysis estimated that the BLR model would correctly predict 86% of an independent test set to pass or fail an on-road assessment. The sensitivity of the model was 91% (29/32 correctly classified as Fail) and the specificity was 78% (14/18 correctly classified as Pass).

Table 4
Proportion of DAVAS referrals who passed or failed the on-road assessment compared with binary logistic regression estimated probability of failing on-road assessment

BLR probability	Outcome of blinded on-road assessment
0.00–0.05	9 pass/0 fail
0.06–0.79	9 pass/6 fail
0.80–1.00	0 pass/26 fail

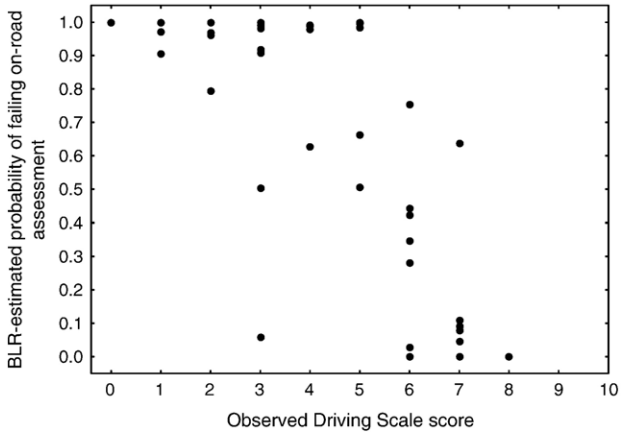

Fig. 3. Correlation between BLR-estimated probability of failing the on-road assessment and observed Driving Scale scores ($r=-0.77$).

Based on *SMCTests* performance, NCRA produced resource demand function curves able to correctly classify 45 of the 50 referrals (i.e., 90% accuracy of training data). The sensitivity of the model was 84% (27/32 correctly classified as Fail) and the specificity was 100% (18/18 correctly classified as Pass). The correlation between the NCRA estimated Driving Scale scores and observed on-road Driving Scale scores was $r=0.87$ (Fig. 4).

Leave-one-out cross-validation analysis estimated that the NCRA model would correctly predict 76% of an independent test set to pass or fail an on-road assessment. The sensitivity of the model was 84% (27/32 correctly classified as Fail) and the specificity was 61% (11/18 correctly classified as Pass).

Neither age nor sex contributed to the accuracy of the BLR or NCRA models.

All 12 control subjects passed the on-road assessment, receiving Driving Scores between 6–9, although the assessor considered one subject marginal. Both the BLR and NCRA models that had been based on data from the 50 referrals predicted on-road passes for 11/12 control subjects (92% specificity for predicting a Pass). Both NCRA and BLR

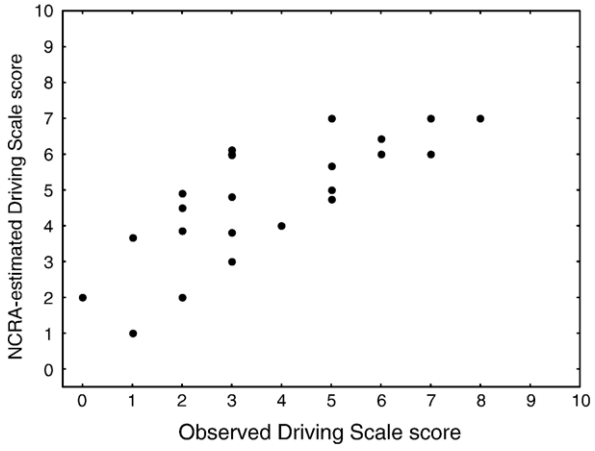

Fig. 4. Correlation between NCRA-estimated Driving Scale scores and observed Driving Scale scores ($r=0.87$).

Table 5
Comparison between observed driving assessment outcome and NCRA- and BLR-predicted driving assessment outcome in an independent set of $n=12$ healthy controls

Control subject	Observed driving assessment outcome/Driving Scale score	NCRA-predicted driving assessment outcome/Driving Scale score	BLR-predicted driving assessment outcome/ probability of Failing on-road
1	Pass/9	Pass/7.0	Pass/0.00
2	Pass/8	Pass/7.0	Pass/0.00
3	Pass/8	Pass/6.7	Pass/0.00
4	Pass/8	Pass/6.4	Pass/0.31
5	Pass/7	Pass/7.0	Pass/0.00
6	Pass/7	Pass/7.0	Pass/0.02
7	Pass/7	Pass/7.0	Pass/0.01
8	Pass/7	Pass/6.6	Pass/0.01
9	Pass/6	Pass/6.8	Pass/0.00
10	Pass/6	Pass/6.7	Pass/0.00
11	Pass/6	Pass/6.6	Pass/0.02
12	Pass/6	Fail/5.0	Fail/0.53

incorrectly predicted a fail for the control who achieved a marginal pass on the on-road assessment. However, both models were marginal in their fail prediction for this subject (NCRA-predicted Driving Scale score 5, BLR probability of 0.53 of failing the on-road assessment). A comparison of the observed driving assessment outcome with NCRA- and BLR-predicted driving assessment outcome is shown in Table 5.

4. Discussion

Our study has shown that, compared with other studies, *SMCTests* have the highest predictive accuracy as estimated in an independent data set and against true on-road driving ability in persons with brain disorders. Furthermore, *SMCTests* are unique in being able to comprehensively and objectively assess both sensory-motor and higher cognitive functions related to driving.

Four previous studies have estimated the predictive accuracy of off-road tests for driving ability in an independent group [14,19,23,32]. Prediction based upon logistic regression of our *SMCTests* data achieved the highest accuracy of 86% with the exception of the NCRA predictive model of Fischer et al. [32] with an accuracy of 89%. However, Fischer et al.'s study was substantially limited by the small number of subjects in their test group ($n=9$) and the on-road assessment being limited to a closed-course circuit. The basic physical and cognitive functions used to build Fischer et al.'s model would likely relate well to the vehicle control aspects assessed on the closed-course driving assessment. Conversely, key higher cognitive function skills needed to respond appropriately to the less predictable and more demanding aspects of true on-road driving, especially relating to intersections, other vehicles, pedestrians, distractions, etc., are not challenged on a closed-course circuit [9,57,58]. Despite these limitations, Fischer et al.'s study provides considerable support for NCRA as a means of predicting performance on a complex high-level task such as driving.

Another study to estimate the predictive accuracy of off-road tests for driving ability in an independent group was based upon the DriveABLE assessment system [21,23]. DriveABLE has a 'high cut-off score', above which an on-road pass in indicated, and a 'low cut-off score', below which an on-road fail is indicated. An on-road assessment is required for all subjects whose scores fall within the 'indeterminate range' between the high and low cut-off scores. A validation study of the predictive accuracy of DriveABLE was undertaken with 431 drivers which included healthy drivers across the age range and patients with a wide variety of medical conditions (e.g., dementia, pulmonary disease, cardiovascular disease, renal disease, head trauma) [39]. The study found that a third of participants received scores in the indeterminate range [23,59], and, of the remaining participants, 94% were correctly predicted to pass or fail the on-road assessment. However, the actual accuracy of pass and fail for all participants was 63% (94% of 67% of subjects correctly classified). Another limitation of DriveABLE is its reliance upon a push-button and touch-screen format which cannot measure or take into account deficits in lower-limb functioning.

Previous studies have also investigated the relationship between off- and on-road tests and determined accuracy in terms of classification models [7–9,15–17,22]. However, these results are based on models which are not validated with separate test data and in which there is a substantial risk of models overfitting the training data. That is, in order to maximize the accuracy of the model, the model becomes too specific to the training data and is unable to generalize well to an independent population. Thus, less weight can be given to the performance of driving assessment systems for which predictive accuracy has been solely determined based on test and training data which are not independent.

Our study has demonstrated that performance on a battery of driving-specific sensory-motor and cognitive tests can be used to accurately predict ability to drive safely and to do so in a range of neurological disorders. Leave-one-out analysis indicated that overall the predictive model based on BLR was more accurate than NCRA at predicting on-road Pass or Fail (BLR 86% vs NCRA 76%). However, NCRA has an important advantage in that it is able to identify which of possibly several deficits a subject may have that most restricts his or her ability to drive safely. This advantage provides an opportunity for focused rehabilitation which might optimally reduce or offset the deficit and enable safe return to the road. The two models provide complementary information, with BLR providing the probability of an on-road assessment pass or fail and NCRA providing information regarding the functional deficit most limiting on-road performance.

While safe driving requires a number of simple component skills, the coordination and organization of these component processes are also important [36]. Thus, driving simulators, which aim to replicate on-road driving, might be considered more useful in determining on-road driving ability than batteries of component function tests. However, driving simulators have, thus far, proven to be less predictive for on-road driving than *SMCTests* [24–27] and are limited by problems with simulator sickness, especially in older drivers [60]. Although *SMCTests* are components based, we believe it is successful at predicting on-road driving ability because (1) on-road driving requires that basic sensory-motor and cognitive component functions are intact and (2) *SMCTests* higher cognitive function tests require integration of the component processes.

Our results (Table 4) show that BLR indicated a high probability (>0.80) of failing the on-road assessment for 81% (26/32) of referrals who did fail the on-road assessment. In contrast, no referrals who passed the on-road assessment had BLR probabilities of 0.80 or higher. Using a BLR probability of 0.80 or higher as the cut-point for determining that an on-road assessment would be too dangerous to undertake, we estimate that up to 80% of potentially dangerous on-road assessments could be avoided each year by DAVAS. The value of this decrease in potentially dangerous on-road assessments should not be underestimated. Despite strong safety measures in place at DAVAS, the use of a true on-road course means that it takes only a few seconds for an unsafe driver to unintentionally drive a vehicle into an oncoming car or hazard. If the new assessment system was established in driving assessment centres throughout New Zealand with its small population (~4 million), we estimate that up to 700 potentially dangerous on-road assessments could be avoided annually.

A limitation of the current study is the relatively small number of referrals and the relatively high proportion of stroke patients. We are working towards obtaining a larger data set to further increase the accuracy and breadth of the predictive models and provide robust verification. Future studies will also increase the number of patients in different diagnostic groups. Despite limitations of the current study, *SMCTests* is unique in being able to comprehensively assess both sensory-motor and higher cognitive functions related to driving. By contrast, other assessment systems either rely on cognitive functions and lack quantitative assessment of sensory-motor functions [9,14,19–23], or rely on assessment of sensory-motor functions with minimal [18] or no tests of higher cognitive functions [32].

Given the promising results achieved in this study with referrals, two additional developments have been initiated. Firstly, a less expensive version of the assessment apparatus has been developed based on portable components (laptop, screen, steering wheel, and pedals). The portable apparatus will offer the full range of *SMCTests* and is designed for driving assessment services provided by specialist driving occupational therapists. Secondly, an abbreviated version of *SMCTests* has also been developed as a screening tool for medical practitioners and is being evaluated through a large group of general practitioners and through a study with healthy older drivers. An accurate screening tool should lead to increased confidence for medical practitioners referring appropriate persons for comprehensive driving assessments

or ratifying the licences of persons who have objectively demonstrated adequate sensory-motor and cognitive abilities needed for safe driving. This will ensure that all persons who are at a potentially increased risk of traffic accident are appropriately identified and referred for assessment.

Accurate off-road estimation of driving ability can minimize the number of on-road assessments of patients who will almost definitely fail, as well as minimizing on-road assessment of patients who are safe to drive. A decrease in unnecessary on-road tests increases cost-effectiveness, assessment efficiency, and, importantly, the safety of referrals, assessors, and other road users by reducing the risk of on-road accidents during on-road assessments.

Acknowledgements

This work was supported through grants from the Foundation for Research, Science, and Technology, Accident Compensation Corporation, and the University of Otago. We thank Associate Professor Christopher Frampton for his statistical advice. We also thank Professor George V. Kondraske and the University of Texas at Arlington for providing the Non-linear Causal Resource Analysis (NCRA) software for our use.

References

[1] Hunt L, Morris JC, Edwards D, Wilson BS. Driving performance in persons with mild senile dementia of the Alzheimer type. J Am Geriatr Soc 1993;41(7):747–52.
[2] Lings S, Jensen PB. Driving after stroke: a controlled laboratory investigation. Int Disabil Stud 1991;13(3):74–82.
[3] Wood JM, Worringham C, Kerr G, Mallon K, Silburn P. Quantitative assessment of driving performance in Parkinson's disease. J Neurol Neurosurg Psychiatry 2005;76(2):176–80.
[4] Hawley CA. Return to driving after head injury. J Neurol Neurosurg Psychiatry 2001;70(6):761–6.
[5] Johnson JE. Urban older adults and the forfeiture of a driver's license. J Gerontol Nurs 1999;25(12):12–8.
[6] (2002). Land Transport Safety Authority *Medical aspects of fitness to drive*. New Zealand.
[7] Fox GK, Bowden SC, Bashford GM, Smith DS. Alzheimer's disease and driving: prediction and assessment of driving performance. J Am Geriatr Soc 1997;45(8):949–53.
[8] Mazer BL, Korner-Bitensky NA, Sofer S. Predicting ability to drive after stroke. Arch Phys Med Rehabil 1998;79(7):743–50.
[9] Galski T, Bruno RL, Ehle HT. Prediction of behind-the-wheel driving performance in patients with cerebral brain damage: a discriminant function analysis. Am J Occup Ther 1993;47(5):391–6.
[10] Sivak M, Olson PL, Kewman DG, Won H, Henson DL. Driving and perceptual/cognitive skills: behavioral consequences of brain damage. Arch Phys Med Rehabil 1981;62(10):476–83.
[11] Schanke AK, Sundet K. Comprehensive driving assessment: neuropsychological testing and on-road evaluation of brain injured patients. Scand J Psychol 2000;41(2):113–21.
[12] Reger MA, Welsh RK, Watson GS, Cholerton B, Baker LD, Craft S. The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. Neuropsychology 2004;18(1):85–93.
[13] De Raedt R, Ponjaert-Kristoffersen I. The relationship between cognitive/neuropsychological factors and car driving performance in older adults. J Am Geriatr Soc 2000;48(12):1664–8.
[14] Nouri FM, Lincoln NB. Validation of a cognitive assessment: predicting driving performance after stroke. Clin Rehabil 1992;6:275–81.
[15] Nouri FM, Tinson DJ, Lincoln NB. Cognitive ability and driving after stroke. Int Disabil Stud 1987;9(3):110–5.
[16] McKenna P, Jefferies L, Dobson A, Frude N. The use of a cognitive battery to predict who will fail an on-road driving test. Br J Clin Psychol 2004;43(Pt 3):325–36.
[17] Brown LB, Stern RA, Cahn-Weiner DA, Rogers B, Messer MA, Lannon MC, et al. Driving Scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. Arch Clin Neuropsychol 2005;20(2):209–15.
[18] Gianutsos R. Driving advisement with the Elemental Driving Simulator (EDS): when less suffices. Behav Res Methods Instrum Comput 1994;26(2):183–6.
[19] Nouri FM, Lincoln NB. Predicting driving performance after stroke. Br Med J 1993;307(6902):482–3.
[20] Ball K, Owsley C. The useful field of view test: a new technique for evaluating age-related declines in visual function. J Am Optom Assoc 1993;64(1):71–9.
[21] Dobbs AR. DriveABLE: a new clinical tool measures driving competency. Can Fam Physician 2000;46:142–3.
[22] Myers RS, Ball KK, Kalina TD, Roth DL, Goode KT. Relation of useful field of view and other screening tests to on-road driving performance. Percept Mot Skills 2000;91(1):279–90.
[23] Dobbs AR. Evaluations for at-risk experienced drivers. Edmonton, Alberta: DriveABLE Testing, Ltd.; 1997.
[24] Lee HC, Cameron D, Lee AH. Assessing the driving performance of older adult drivers: on-road versus simulated driving. Accid Anal Prev 2003;35(5):797–803.
[25] Wald JL, Liu L, Reil S. Concurrent validity of a virtual reality driving assessment for persons with brain injury. Cyberpsychol Behav 2000;3 (4):643–53.
[26] Lew HL, Poole JH, Lee EH, Jaffe DL, Huang H-C, Brodd E. Predictive validity of driving-simulator assessments following traumatic brain injury: a preliminary study. Brain Inj 2005;19(3):177–88.
[27] Lundqvist A, Gerdle B, Ronnberg J. Neuropsychological aspects of driving after a stroke — in the simulator and on the road. Appl Cogn Psychol 2000;14(2):135–50.
[28] McKnight AJ, McKnight AS. Multivariate analysis of age-related driver ability and performance deficits. Accid Anal Prev 1999;31 (5):445–54.
[29] Marottoli RA, Richardson ED, Stowe MH, Miller EG, Brass LM, Cooney Jr LM, et al. Development of a test battery to identify older drivers at risk for self-reported adverse driving events. J Am Geriatr Soc 1998;46(5):562–8.
[30] Sims RV, McGwin Jr G, Allman RM, Ball K, Owsley C. Exploratory study of incident vehicle crashes among older drivers. J Gerontol A Biol Sci Med Sci 2000;55(1):M22–7.
[31] Lee HC, Lee AH, Cameron D, Li-Tsang C. Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. J Safety Res 2003;34(4 Suppl):453–9.
[32] Fischer CA, Kondraske GV, Stewart RM. Prediction of driving performance using nonlinear causal resource analysis. Proc 2nd Joint Eng Med Biol Soc/Biomed Eng Soc Conf, vol. 2; 2002. p. 2473–4.
[33] Hoffman JD, Brown TL, Lee JD, McGehee DV. Comparisons of braking in a high fidelity simulator to braking on a test track. Trans Res Rec 2002;1803:59–65.
[34] Rizzo M, Reinach S, McGehee D, Dawson J. Simulated car crashes and crash predictors in drivers with Alzheimer disease. Arch Neurol 1997;54:545–51.
[35] Rizzo M, McGehee DV, Dawson JD, Anderson SN. Simulated car crashes at intersections in drivers with Alzheimer disease. Alzheimer Dis Assoc Disord 2001;15(1):10–20.
[36] Graydon FX, Young R, Benton MD, Genik II RJ, Posse S, Hsieh L, et al. Visual event detection during simulated driving: Identifying the neural correlates with functional neuroimaging. Transp Res Part F Traffic Psychol Behav 2004;7(4–5):271–86.

[37] Uc EY, Rizzo M, Anderson SW, Shi Q, Dawson JD. Unsafe rear-end collision avoidance in Alzheimer's disease. J Neurol Sci 2006;251 (1–2):35–43.
[38] Stolwyk RJ, Charlton JL, Triggs TJ, Iansek R, Bradshaw JL. Neuropsychological function and driving ability in people with Parkinson's Disease. J Clin Exp Neuropsychol 2006;28(6):898–913.
[39] Dobbs AR. The development of a scientifically based driving assessment and standardization procedures for evaluating medically at-risk drivers. Canadian Multidisciplinary Road Safety Conference XV, New Brunswick, Canada; 2005.
[40] Jones RD. Measurement of sensory-motor control performance capacities: tracking tasks. In: Bronzino JD, editor. The biomedical engineering handbook — biomedical engineering fundamentals, vol. 1. Boca Raton, Florida: CRC Press; 2006. p. 1–25. [77].
[41] Jones RD, Sharman NB, Watson RW, Muir SR. A PC-based battery of tests for quantitative assessment of upper-limb sensory-motor function in brain disorders. Proc 15th Ann Int Conf IEEE Eng Med Biol Soc, vol. 15; 1993. p. 1414–5.
[42] Jones R, Giddens H, Croft D. Assessment and training of brain-damaged drivers. Am J Occup Ther 1983;37:754–60.
[43] Croft D, Jones RD. The value of off-road tests in the assessment of driving potential of unlicensed disabled people. Br J Occup Ther 1987;50:357–61.
[44] Jones RD, Donaldson IM, Parkin PJ. Impairment and recovery of ipsilateral sensory-motor function following unilateral cerebral infarction. Brain 1989;112:113–32.
[45] Jones RD, Donaldson IM, Parkin PJ, Coppage SA. Impairment and recovery profiles of sensory-motor function following stroke: single-case graphical analysis techniques. Int Disabil Stud 1990;12:141–8.
[46] Jones RD, Donaldson IM. Fractionation of visuoperceptual dysfunction in Parkinson's disease. J Neurol Sci 1995;131:43–50.
[47] Jones RD, Donaldson IM, Timmings PL. Impairment of high-contrast visual acuity in Parkinson's disease. Mov Disord 1992;7:232–8.
[48] Jones RD, Donaldson IM, Sharman NB. A technique for the removal of the visuospatial component from tracking performance and its application to Parkinson's disease. IEEE Trans Biomed Eng 1996;43:1001–10.
[49] Dalrymple-Alford JC, Kalders AS, Jones RD, Watson RW. A central executive deficit in patients with Parkinson's disease. J Neurol Neurosurg Psychiatry 1994;57:360–7.
[50] Jones RD, White AJ, Lawson KH, Anderson TJ. Visuoperceptual and visuomotor deficits in developmental stutterers: An exploratory study. Hum Mov Sci 2002;21:603–19.
[51] Dalrymple-Alford JC, Kerr PA, Jones RD. The effects of alcohol on driving-related sensorimotor performance across four times of day. J Stud Alcohol 2003;64:93–7.
[52] Heitger MH, Anderson TJ, Jones RD, Dalrymple-Alford JC, Frampton CM, Ardagh MW. Eye movement and visuomotor arm movement deficits following mild closed head injury. Brain 2004;127(3):575–90.
[53] Hopkins WG. A new view of statistics, vol. 2004. Internet Society for Sport Science; 2000.
[54] Dawson B, Trapp RG. Statistical methods for multiple variables. Basic & clinical biostatistics. McGraw-Hill Professional; 2001.
[55] Kondraske GV. The elemental resource model for human performance. In: Bronzino JD, editor. The Biomedical Engineering Handbook — Biomedical Engineering Fundamentals, vol. 1. Boca Raton, Florida: CRC Press; 2006. p. 1–19. [75].
[56] Witton IH, Frank E. Credibility: evaluating what's been learned. Data Mining. San Francisco, CA: Morgan Kaufmann; 1999. p. 119–56.
[57] Fox GK, Bowden SC, Smith DS. On-road assessment of driving competence after brain impairment: review of current practice and recommendations for a standardized examination. Arch Phys Med Rehabil 1998;79(10):1288–96.
[58] Galski T, Ehle HT, Bruno RL. An assessment of measures to predict the outcome of driving evaluations in patients with cerebral damage. Am J Occup Ther 1990;44(8):709–13.
[59] (1999). National Highway Traffic Safety Administration *Safe mobility for older people notebook* (DOT HS 808853): US Department of Transportation.
[60] Liu L, Watson B, Miyazaki M. VR for the elderly: quantitative and qualitative differences in performance with a driving simulator. Cyberpsychol Behav 1999;2(5):1–35.