AN EVALUATION OF DRIVING ABILITY IN A SIMULATOR: A GOOD PREDICTOR OF DRIVING ABILITY AFTER STROKE?

Ann-Helen Patomella, Anders Kottorp
Karolinska Institutet
Stockholm, Sweden
E-mail: ann-helen.patomella@neurotec.ki.se

Summary: The aim of this pilot-study was to evaluate the predictive quality of a driving evaluation in a simulator for people with stroke. Twenty-seven participants with stroke completed a driving evaluation in a technically advanced driving simulator. The participants also completed an on-road driving evaluation with a pass or fail outcome. Measures generated from the assessment tool, P-Drive, were compared to the on-road driving evaluation, in order to determine a definitive cut-off score. The definitive cut-off score of P-Drive was determined to 1,5 logits, however this score did not show satisfactory sensitivity. Instead of a definitive cut-off score, a cut-off zone of 1,4 to 1,7 logits was recommended. Six drivers within the cut-off zone were recommended for further evaluation of their driving ability.

INTRODUCTION

After a brain injury (e.g., stroke) there is a need to investigate people's ability to continue driving, but there are risks involved in assessing people with an acquired brain injury in an on-road evaluation (Rizzo, McGehee, Dawson, & Anderson, 2001). On-road assessments, however, have been criticized for their lack of offering challenging situations and limited standardization procedures (Christie, 1996). A driving simulator could therefore be an alternative assessment context for assessing driving ability that is both safe and could also be standardized and challenging. No assessment tool has so far been developed for assessing driving ability in a simulator. Thus, there is a lack of valid and reliable assessment tools for assessing driving ability in a practical setting that tells whether a driver with stroke could continue driving.

A validation study of a new assessment tool, P-Drive (Performance Analysis of Driving Ability), for evaluating driving ability as performed in a simulator, has been conducted with people with stroke (Patomella, Caneman, Kottorp, & Tham, 2004). The results supported evidence of internal scale validity as well as validity in relation to response processes. The items in the scale measured a single construct, indicating unidimensionality. The results also demonstrated that the participants represented a wide range of driving ability, suggesting that P-drive may be used for people with different levels of driving ability. The hierarchy of the items was in line with assumptions made in previous research in the field of driving, supporting construct validity. Altogether the results were promising, but there was still a clinical need to evaluate the use of P-Drive in order to predict whether a person is able or unable to drive in real traffic.

In order for the examiner to predict an outcome (e.g., pass /fail; able /unable) an evaluation must include a cut-off criteria (Streiner & Norman, 2003). Driving evaluations in simulators have earlier been criticized for problems in interpretation of the outcomes, due to the absence of reliable norms or cut-off criteria (Galski, Ehle, McDonald, & Mackevich, 2000). The present

study aimed to determine and evaluate the predictive quality of P-Drive when used in a simulator. The following research question was addressed:

- Can P-Drive separate between able and unable drivers, i.e., predict those who are able or unable to drive after stroke?

METHODS

Participants

The present study was approved by the Ethical Committee of Northern Stockholm. The sample consisted of a consecutive series of 27 participants with stroke. The participants were invited from the Stockholm Stroke Registry (*Welcome to the National Stroke Register*, 2005). The inclusion criteria for the participants were that time since onset of stroke had to be between 3 and 12 months and age had to be between 30 and 70 years. The participants´ characteristics are further described in Table 1.

Table 1. Clinical characteristics of the participants

Participants (n=27)	Mean age (years)	Time since onset (months)
Male (n=24):	M: 57,5	M: 7,5
Female (n=3):	SD: 8,0	SD: 3,8

Procedure

The evaluation started with a short interview of medical and driving history. All of the participants got the chance to practise in the simulator in a special practise programme that allowed the drivers to get used to the controls of the car, like pedals, gear and steering. The participant decided when he/she did not need any more practise (commonly after 15-30 minutes). Thereafter, the participants underwent driving evaluation in the test programme in the simulator. An occupational therapist with previous experience of performing driving evaluations scored each participant's driving performance using P-Drive (Performance Analysis of Driving Ability). The scoring took place directly after the test drive. All of the participants also completed an evaluation on-road with a traffic inspector from the National Road Safety Administration. The on-road evaluation was conducted on a separate occasion, within a couple of months of the evaluation in the simulator.

Instrumentation

Simulator driving. An interactive, realistic and technically advanced simulator, situated at the Karolinska University Hospital (Stockholm, Sweden), was used in this study. The simulator was developed and used by clinical occupational therapists to assess driving ability after stroke. The simulator consisted of a real, but truncated car (see Figure 1). As in real driving, the driver had to operate the pedals, turn the steering-wheel, etc. The driving program was projected onto three big screens with a field-of-view of 135 degrees. An audio system generated synchronised natural sounds, such as the engine running and radio messages. The test-program took approximately 45-60 min, depending on the speed of the car and the number of mistakes made by the driver.

The test-program consisted of 70 traffic situations, each presenting a different challenge to the driver. The situations required the driver to deal with situations that one would frequently encounter when driving a car, and had been designed to specifically address situations assumed to be extra challenging for persons with acquired brain injury.

Figure 1. Argus Driving simulator

P-Drive (Performance Analysis of Driving Ability). The evaluation of driving performance in the simulator was carried out by direct observation of the driving in the test-programme. During the driving in the simulator, each traffic situation that occurred during the test involved a number of actions that the driver had to perform. The observer then judged the quality of performance on 20 items that were defined in P-Drive. During a complete simulator driving in the test-programme, all of the items in P-Drive were observable several times and in different challenging situations. The test-leader assigned the score to an item according to the most incompetent performance observed during the session (for example, see Table 2). P-Drive consists of a test sheet (A4) where the 20 items are scored using a 4-point criterion-referenced rating scale based on the quality of performance: 4 = competent performance, 3 = questionable, 2 = ineffective, and 1 = incompetent.

Table 2. Examples of traffic behaviour and items scored

Traffic situation/behaviour	Items (actions) scored in P-Drive
Participant collided with a car coming from the right (had to give right-of-way).	Paying attention to the right, (score 1) Giving right-of-way, (score 1) Controlling speed, (scored depending on the way the participant adapted the speed for the oncomming crossing).
Participant stopping at stop-sign.	Obeying stop-sign, (score 4) Using pedals, (score 4)

On-road evaluation. The on-road test was conducted by an inspector from the Swedish National Road Administration. The inspector had several years of experience in evaluating drivers with a neurological disability. The assessment consisted of a set route in the Stockholm area and took approximately 70 min to complete. The route included a diversity of road types (suburban, central streets, motorways and country roads). The outcome of the driving was dichotomous as either pass or fail.

Data analysis

A computer application of Rasch analysis, FACETS (LINACRE, 2005), was used to generate equal interval measures of driving ability (logits) based on raw score data. In this study, the use of an external outcome (e.g., the result of the on-road driving evaluation) was set as a criterion for establishing a cut-off score for P-Drive. A Receiver Operation Characteristics (ROC) curve was calculated in order to determine a potential cut-off score of P-Drive. In this study, the cut-off score was selected due to the highest achieved sensitivity in combination with highest achieved specificity. The accuracy of a test is judged qualitatively by the position of its ROC curve, and there is a balancing between the level of sensitivity and specificity when choosing a cut-off (Chu, 1999). Based on the estimated cut-off score from the ROC-curve, sensitivity, specificity and the predictive values of P-Drive were calculated. In addition to the cut-off score, a cut-off zone was also estimated based on the descriptive inspection of the P-Drive scores compared to the on-road driving evaluation.

RESULTS

Definitive cut-off; 1,5

The use of a ROC-curve determined cut-off score to 1,5 logits. The sensitivity of P-Drive cut-off score was 70% and the specificity was 81% (see Table 3). The positive predictive value of P-Drive cut-off score was estimated to 67%.

Table 3. Sensitivity and specificity of P-Drive using a cut-off score of 1,5 logits

Sensitivity Probability that P-Drive was positive when the client passed on-road evaluation	0,70	**Specificity** Probability that P-Drive was negative when the client failed on-road eval	0,81
Positive predictive value (PV+) The percentage of people with P-Drive results over cut-off score who actually	0,67	**Negative predictive value (PV-)** The percentage of people with P-Drive results under cut-off score who actually failed on-road evaluation	0,81

Cut-off zone 1,4-1,7 logits

The cut-off range area was set to be between 1,4 to 1,7 logits and was based on clinical judgement and a visual inspection of the data (see Figure 2). Within the cut-off zone were six participants, in which 3 passed and 3 failed (i.e., the participants had a 50/50 chance of passing the driving evaluation in the simulator).

Table 4. Sensitivity and Specificity of P-Drive using a cut-off zone of 1,4-1,7 logits

Sensitivity	0,80	Specificity	0,86
Positive predictive value (PV+)	0,80	Negative predictive value (PV-)	0,86

The sensitivity of P-Drive was 80%, which is the amount correctly classified as having acceptable driving ability using P-Drive cut-off zone 1,4-1,7. The specificity of P-Drive cut-off zone was 86%, which was the amount correctly classified as having unacceptable driving ability by P-Drive. The PV+ also increased from 68% to 80% by the use of a cut-off zone.
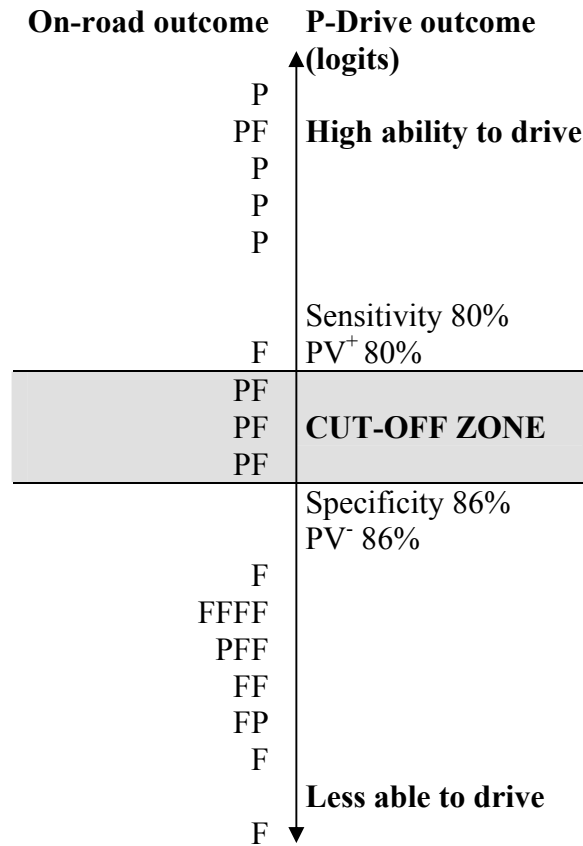
On-road outcome   **P-Drive outcome**
(logits)

P
PF   **High ability to drive**
P
P
P

Sensitivity 80%
F PV^+ 80%
PF
PF   **CUT-OFF ZONE**
PF
Specificity 86%
PV^- 86%
F
FFFF
PFF
FF
FP
F
Less able to drive
F

Figure 2. Distribution of the participants along the logits scale of P-Drive
F = failed the on-road evaluation, P = passed the on-road evaluation

DISCUSSION

The aim of the study was to determine and investigate the predictive qualities of P-Drive. By determining a cut-off score at 1,5 logits, the sensitivity was lower than expected, indicating that the P-Drive cut-off score failed to correctly identify drivers that passed the on-road evaluation. This result may be due to the fact that the number of participants was too small and may not be enough to determine a precise cut-off score. Another potential reason could be a lack of transfer-effects from a drive in a simulator to real traffic. Finally, the on-road evaluations were conducted

within a couple of months after the simulator drives, and spontaneous recoveries may therefore also have occurred among the participants, resulting in a lower sensitivity than expected.

Before setting a definitive cut-off score, decision-makers must consider the consequences of incorrect judgements and the prevalence of the problem in the population being tested (Swets, Dawes, & Monahan, 2000). People with stroke often have decreased cognitive capacities as attention and speed of information processing (Lundqvist, Gerdle, & Rönnberg, 2000), which could affect their driving ability and therefore need to be detected. Achieving both a high sensitivity and a high specificity is of importance in assessment tools for assessing driving after stroke, but may never be fulfilled completely. The choice of a definitive cut-off score of a test of driving ability must therefore be carefully selected and be judged on various aspects in relation to these concepts. People who seem most likely to endanger others in traffic need to be identified and prevented from driving, for both society's and their own safety. At the same time, interfering in the lives of people who do not need to be prevented from driving independently may be very trying both for the individual and society. Thus, there was a need to improve the predictive qualities of P-Drive when using a cut-off at 1,5 logits.

As an alternative to the definitive cut-off score, a cut-off zone was therefore considered as a more clinically valid outcome ruler. By the use of such area the sensitivity increased (80%), but also regained an acceptable level of specificity. Such a cut-off zone may be clinically useful for professionals, since it gives information about when to pass or fail a driver with stroke, but also an indication of when the ability to drive needs to be further evaluated with either neuropsychological testing or with another practical evaluation (e.g., a simulator or on-road).

The on-road driving evaluation is usually used as a criterion for standardizing assessment tools for assessing driving ability (Christie, 1996). However, the on-road evaluation has been questioned for its inability to pick up unsafe driving behaviour (Christie, 1996; Vägverket, 2001), and also for its risky nature due to the fact that an unsafe patient may be evaluated (Korner-Bitensky, Sofer, Gelinas, & Mazer, 1998; Vägverket, 2001). A normal on-road driving test is seen as a rough estimate of a participant's ability and it is unusual that the driving is challenging enough to pick up risky behaviour due to cognitive impairment (Vägverket, 2001). It may be impossible to identify all unable drivers after stroke, but there is an assessment tool (P-Drive) that can help clinicians to identify able and unable drivers, as well as those who need further evaluation. The three participants in the P-Drive cut-off zone that passed the on-road evaluation were permitted to continue driving by their physician, while the three participants within the P-Drive cut-off zone that failed on-road where called for reassessment in the simulator (6-12 months from the on-road evaluation). However, they all passed the criterion the second time. These clinical results also suggest that driving ability may not be a static ability within a person, but may vary and develop over time. Participants in the cut-off zone might not be ready to resume driving at the time for testing them, but instead need more time to recover from their stroke. The results of this study indicate the use of the P-Drive cut-off zone as a predictive assessment tool for use in the simulator when assessing real driving ability after stroke.

CONCLUSION

A definitive cut-off score on P-Drive to predict actual driving ability resulted in acceptable specificity, but unacceptable low sensitivity. A cut-off zone was instead proposed, resulting in

both a higher sensitivity and specificity. Thus, the drivers within the cut-off zone are within a borderline of pass and fail and it is therefore suggested that they be evaluated further. Further tests could include retesting in the simulator, or other tests related to detect capacities and abilities necessary for driving ability.

ACKNOWLEDGMENTS

REFERENCES

Christie, N. (1996). Assessing Driving Fitness Following Brain Injury or Illness: A research review. *TRL Report 208*, 1-21.

Chu, K. (1999). An introduction to sensitivity, specificity, predicitive values and likelihood ratios. *Emergency Medicine, 11*, 175-181.

Galski, T., Ehle, H. T., McDonald, M. A., & Mackevich, J. (2000). Evaluating fitness to drive after cerebral injury: basic issues and recommendations for medical and legal communities. *Journal of Head Trauma Rehabilitation, 15*(3), 895-908.

Korner-Bitensky, N. A., Sofer, S., Gelinas, I., & Mazer, B. L. (1998). Brief or new. Evaluating driving potential in persons with stroke: a survey of occupational therapy practices. *American Journal of Occupational Therapy, 52*(10), 916-919.

Linacre, J. (2005). Facets: Many-facet Rasch Measurement Programme (Version 3.56.1). Chicago: MESA Press.

Lundqvist, A., Gerdle, B., & Rönnberg, J. (2000). Neuropsychological aspects of driving after stroke: in the simulator and on the road. *Applied Cognitive Psychology, 14*, 135-150.

Patomella, A., Caneman, G., Kottorp, A., & Tham, K. (2004). Identifying scale and person response validity of a new assessment of driving ability. *Scandinavian Journal of Occupational Therapy, 11*(2), 70-77.

Rizzo, M., McGehee, D. V., Dawson, J. D., & Anderson, S. N. (2001). Simulated Car Crashes at Intersections in Drivers With Alzheimer Disease. *Alzheimer Disease and Associated Disorders, 15*(1), 10-20.

Streiner, D., & Norman, G. (2003). *Health Measurement Scales- a practical guide to their development* (Third ed.). New York: Oxford University Press.

Swets, J., Dawes, R., & Monahan, J. (2000). Better decisions through science. *Scientific American* (October), 82-87.

Welcome to the National Stroke Register. (2005). Retrieved 2005-03-14

Vägverket. (2001). *Trafikmedicin*. Borlänge: Vägverket, Trafikmedicinska rådet.