How effective is the Trail Making Test (Parts A and B) in identifying cognitively impaired drivers?

BONNIE M. DOBBS[1], SIMRAN S. SHERGILL[2]

[1]Family Medicine, University of Alberta, 205 8215 112 Street, Edmonton, Alberta, Canada T6G 2C8
[2]Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

Address correspondence to: B. M. Dobbs. Tel: 1 780 492 0374; Fax: 1 780 492 8191; Email: bdobbs@ualberta.ca

Abstract

Background: the medical community plays an important role in identifying drivers who may no longer be competent to drive due to illnesses such as dementia. Several office-based cognitive screening tools are currently used by the medical community, e.g. Mini-Mental State Examination, Trail Making Test (TMT), to assist in the identification of cognitively impaired (CI) at-risk drivers. However, the predictive validity of these tools is questionable.
Objective: to examine the predictive power of the TMT for on-road driving performance.
Methods: data from a prospective sample of CI and healthy older drivers were collected. TMT-A and -B (time and errors) served as predictor variables, with pass/fail on a scientifically based on-road assessment used as the criterion variable. Receiver operating characteristic (ROC) curve analysis was used to assess overall 'diagnostic' accuracy of TMT-A and -B for driving competency. Cut points from previous studies/guidelines were used to assess predictive power.
Findings: a total of 134 older drivers (mean age = 75.30; SD = 7.83) participated: 87 healthy controls and 47 CI individuals. All predictor variables, with the exception of TMT-A errors, were significantly correlated with driving outcome. However, results from ROC curve analyses indicated that only TMT-A and -B total time had moderate discriminative abilities. Results also indicate that the power of the TMT is the lowest where physicians need it most (e.g. identifying CI patients whose driving skills have declined to an unsafe level).
Conclusion: TMT-A and -B outcomes are most likely to be inaccurate in those whose driving competency has declined to an unsafe level, resulting in risks to both individual and public safety.

Keywords: *automobile driving, mild cognitive impairment, dementia, screening tools, primary health care, older people*

Introduction

There now is a substantial body of literature documenting the relationship between medical conditions and impaired driving performance [1, 2–4]. These medical conditions can occur at any age but the likelihood of their occurrence increases significantly with age [3, 5]. For older drivers, driving impairments are most likely the result of changes in functional abilities due to age-associated medical conditions and/or medications used to treat those medical conditions [6, 7]. Illnesses that affect cognitive functioning are of concern given the impact that declines in cognition have on driving performance [2–4, 8]. Although some illnesses that affect cognitive functions necessary for driving are obvious (e.g. dementia and stroke), others are less obvious (e.g. chronic renal failure [9], 10, 11], congestive heart failure [12, 13], diabetes [14, 15] and depression [16, 17]).

Studies using on-road assessments indicate that drivers with dementia show declines in driving performance compared with same-aged individuals without dementia [18, 19, 20, 21, 22, 23–26]. When using crash rates as an outcome measure, drivers with dementia also have higher crash rates [2, 4, 27–29, 30–32, 33 but see 34] with at-fault crash rates that are 3.3 times (95% CI: 1.84–5.59) higher than age-, gender- and place of driving-matched controls [2].

A number of screening tools used to identify cognitive impairment (CI) [see 35] often are used in the clinical setting for assisting with in-office assessment of fitness to drive [36]. Those tools include the Mini-Mental State Examination (MMSE) [37] and the Trail Making Test (TMT) Parts A and B (TMT-A and -B) [38]. Although the MMSE and the TMT-A and -B were not designed as screening tools for identifying declines in *driving*, they often are used or recommended for use in that capacity. There is increasing evidence

that the MMSE is insensitive for detecting drivers with a dementia whose driving abilities have declined to an unsafe level [22, 39]. However, there has been less research on the utility of the TMT for decisions about driving.

Two recent reviews on the use of TMT-A and -B in assessments of driving competency [see 36, 40] indicate that the vast majority of studies have simply investigated the *relationship* between TMT-A and -B scores and driving outcomes. In general, results indicate either no relationship [e.g. 41, 42] or low-to-moderate correlations in the range of 0.16–0.52 [e.g. 25, 43]. Notably, few studies have employed cut points to assess the accuracy of TMT-A and -B for predictions of driving competency [42, 44–49]. Despite the limited findings, results from TMT-A and -B often are used in clinical settings for decisions about driving, with recommendations for their use from both the American and Canadian Medical Associations [50, 51].

The primary objective of this research was to examine the accuracy of established cut points from previously published driving research, medical fitness to drive guidelines and recent normative data on TMT-A and -B for identifying declines in driving competency in a sample of cognitively intact and CI drivers.

Methods

Using prospective methodology, 134 participants aged 65 and older were recruited into the study, with the sample consisting of 47 individuals with CI and 87 healthy, community dwelling seniors (HC). Inclusionary criteria were: (i) clinically diagnosed with CI with or without dementia (for the CI group); (ii) holding a valid driver's licence and actively driving on entry to the study (defined as driving at least one day a week) and (iii) fluent in English. The study protocol received ethical approval by our University's Health Research Ethics Board. Demographic information and a medical and driving history were obtained from all participants, followed by the standardised administration of the TMT-A and -B and the MMSE [37]. Participants then completed an on-road driving assessment [20], given by an experienced driving evaluator on a standardised road course, with the examiner blinded to medical and driving history and cognitive test outcomes.

Predictor measures

The TMT is a test of visual conceptual and visuo-motor tracking. TMT-A purportedly measures attention, visual search and motor function, whereas TMT-B is seen as a measure of executive functioning, speed of attention, visual search and motor function [52]. Outcome measures for both tasks include time to completion and number of errors [53].

Analyses

Baseline characteristics and group differences between the HC and the CI group were compared using MANOVA/ Pearson correlations; Pearson correlational analyses were used to examine the relationship between performance on TMT-A and -B (time and errors) and on-road outcome. Level of significance was set at an alpha of 0.05, using two-tailed tests. Receiver operating characteristic (ROC) curve analysis was used to assess the overall 'diagnostic' accuracy of TMT-A and -B for driving competency using area under the curve (AUC).

Results

Demographics and cognitive test outcomes

The demographics and cognitive test outcomes for the sample as a whole and as a function of cognitive status are presented in Supplementary data available in *Age and Ageing* online, Appendix 1 Table SA. Small-to-moderate correlations were obtained between TMT-A and -B outcomes and on-road outcome [TMT-A time ($r = -0.24$, $P < 0.01$); TMT-B time ($r = -0.42$, $P < 0.01$); and TMT-B errors ($r = -0.24$, $P < 0.01$), with a non-significant relationship for TMT-A errors ($r = -0.14$, $P > 0.05$)].

Predictive properties of TMT-A and -B

The results of the ROC analyses are provided in Figure 1, with AUC values provided in Supplementary data available in *Age and Ageing* online, Appendix 1 Table SB. As shown in Figure 1, the AUC values for TMT-A and -B outcomes ranged from 0.56 to 0.83. As can be seen from the lower and upper limits of the 95% CI, the intervals for each of the outcome measures are relatively wide, indicating a lower precision of the estimate.

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for previously reported cut points and from cut points that represent the best balance between sensitivity and specificity are shown in Table 1. As can be seen, sensitivity was low across all three measures (TMT-A time and TMT-B time and errors), with relatively high specificities. Finally, ROC curve analyses allowed for the identification of cut points that provided the best balance between sensitivity and specificity. The identified cut points resulted in a combination of sensitivity/specificity values that ranged from 73%/68% to 81%/51%, with no cut point for TMT-A (errors) providing a 'good' balance between sensitivity and specificity. Overall, the PPVs (probability of a decline in driving competency) given a positive test result (failed the on-road assessment) ranged from 27% to 77% across TMT-outcomes. The NPVs (probability of no decline in driving competency) given a negative test result (passed the on-road assessment) ranged from 62% to 93% across TMT-outcomes.
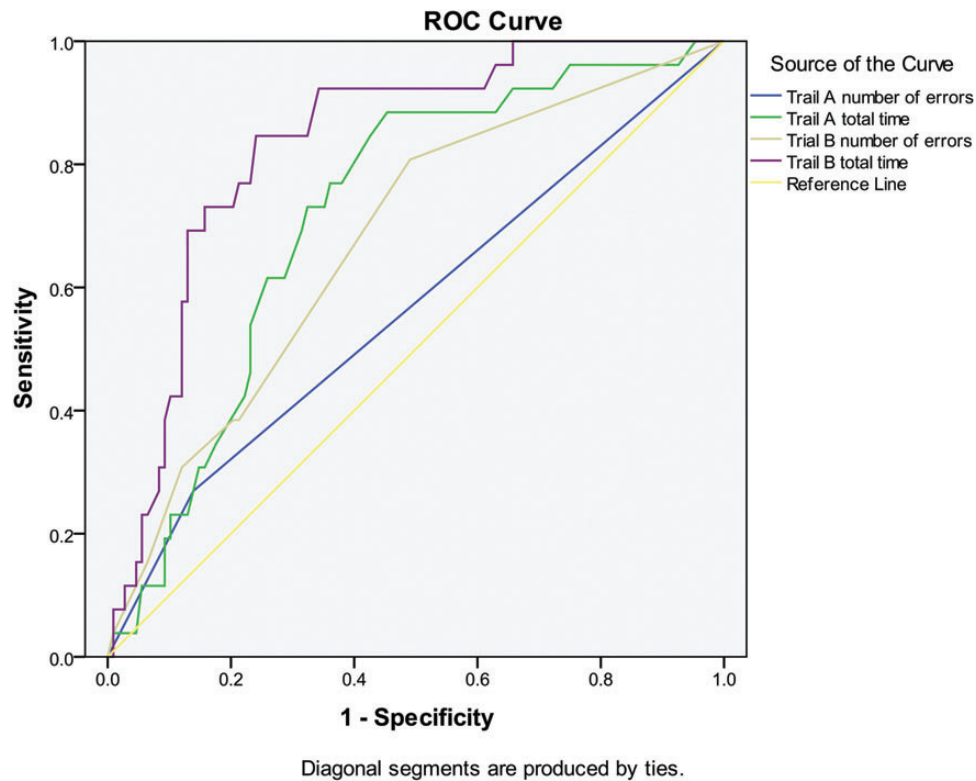
Diagonal segments are produced by ties.

Figure 1. ROC curves of TMT-A and -B (time and errors) for predictions of driving outcome (pass/fail) ($n = 134$).

Table 1. Sensitivity and specificity of TMT-A and -B (based on cut points from driving studies, medical guidelines, normative data and those that maximise sensitivity and specificity from the ROC curve analyses)

Cut points from driving studies/medical guidelines	Pass	Fail	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
TMT-A time	≤39.5 s	>39.5 s	77	62	77	62
TMT-A errors[a]	0	≥1	27	86	32	83
TMT-B time	<180 s	≥180 s	50	88	50	88
TMT-B errors	<3	≥3	31	88	38	84
Cut points from normative data						
TMT-A time	≤68 s	>68 s	12	93	27	81
TMT-A errors[b]	≤1	>1	—	—	—	—
TMT-B time	≤191	> 191	39	91	50	86
TMT-B errors	≤2	>2	31	88	38	84
Cut points that maximise sensitivity and specificity						
TMT-A time	≤42.50	>42.50	73	68	35	91
TMT-A errors[c]	—	—	—	—	—	—
TMT-B time	≤140.50	>140.50	77	77	44	93
TMT-B errors	≤1	>1	38	79	30	84

PPV, positive predictive value; NPV, negative predictive value.
[a]No participants made more than one error, thus the cut points of 0/>1 were used as opposed to ≤1 (pass) and ≥2 (fail) used in the Bédard *et al.* study [54].
[b]No participants made more than one error on TMT-A, precluding analyses on the normative cut point of >1 error.
[c]There was no cut point on TMT-A (errors) that provided a good balance between sensitivity and specificity.

Discussion

The current findings showed significant correlations between three of the four TMT-outcomes and outcomes from the on-road assessment, findings consistent with previous research [e.g. 54–56]. However, despite these significant *correlations*, results from ROC curve analyses indicate that only TMT-A and -B total time have moderate discriminative abilities, but the confidence intervals for all four tests are relatively wide. These results support the assertion from Mathias and Lucas [36] that knowledge of *associations* (e.g. correlations) between cognitive tests and driving performance are of little use for clinical decision making.

These same authors [36] also recommend researchers in this area move away from examining the association between

cognitive tests and driving outcomes, and move to the use of cut points that would allow for an examination of the predictive properties (e.g. sensitivity, specificity, PPV, NPV) of the test (s) under investigation in relation to driving outcomes. Results from this study are revealing in that previously reported and recommended cut points are most effective in identifying drivers *who are competent to drive* (e.g. specificity) but are *ineffective* in identifying drivers *who are no longer competent to drive* (e.g. sensitivity). Results also indicate that cut points selected to balance sensitivity and specificity most often increase the sensitivity of the test with an expected reduction in specificity.

In our analyses, we did not adjust for age and years of education. Recent evidence indicates that adjustments for age and education *diminish* the predictive accuracy of neuropsychological measures for real-world activities such as driving [57]. Those authors report close to a 50% loss of power in predicting driving ability with neuropsychological tests after controlling for age and education. As noted by the authors, although demographic adjustments are an important feature of neuropsychological assessment, when 'the issue at hand is an individual examinee's ability to perform a real-world activity with universal demands for adequate levels of cognitive abilities from all, regardless of demographic characteristics, it is the examinee's absolute levels of relevant ability that are pertinent' (p. 685). Similar arguments have been made by other authors [58, 59].

The present findings have important implications for clinical assessments of fitness to drive for CI patients. While identifying CI patients who remain 'fit-to-drive' is important, detecting drivers with CI who are no longer safe to drive is critical from a public safety perspective. It also is the case that, although driving situations differ in terms of the competence required [60], the physician's role is to judge the patient's competence relevant to driving situations commonly encountered and accommodated by competent drivers. Emerging evidence indicates that there is a need to shift away from the use of traditional neuropsychological tests that are designed to identify CI to tests that have been developed specifically for decisions about driving, and that are suitable for use in the clinical setting. Ideally, those tests will capture the 'interaction of multiple cognitive domains' which is what is needed for a complex activity such as driving.

Key points

- The medical community plays an important role in identifying unsafe drivers.
- All drivers with a progressive dementia will become unsafe to drive at some point in their illness.
- Screening tools for CI may not be valid for decisions about driving.
- Using previously recommended cut points, sensitivity and specificity were low for all TMT-outcomes.
- Results from this research indicate that the power of the TMT is lowest where physicians need it most.

Conflicts of interest

The outcome (road test) used in driving research is part of the DriveABLE assessment. DriveABLE is not identified in the paper but we do reference a published paper on the assessment (Reference 20—Dobbs A, Heller RB, Schopflocher D. A comparative approach to identify unsafe older driver. *Accid Anal Prev.* 1998 May; 30 (3):363–70. Dr Allen Dobbs is the spouse of the first author (B.D.) and is Chief Scientific Officer of DriveABLE Assessment Centres—a University of Alberta spin off company. As noted above, DriveABLE is not identified in the paper and the paper is not oriented towards the on-road assessment.

Funding

Supplementary data

Supplementary data mentioned in the text is available to subscribers in *Age and Ageing* online.

References

The very long list of references supporting this review has meant that only the most important are listed here and are represented by bold type throughout the text. The full list of references is available on the Supplementary data in *Age and Ageing* online, Appendix 1.

2. Diller E, Cook L, Leonard D, Dean JM, Reading J, Vernon D. Evaluating drivers licensed with medical conditions in Utah, 1992–1996. Washington, DC: National Highway Traffic Safety Administration, 1999. Report No.: DOT HS 809 023.
3. Dobbs BM. Medical conditions and driving: a review of the literature (1960–2000). Washington, DC: U.S. Department of Transportation, 2005. Report No. DOT HS 809 690.
4. Vaa T. Impairment, diseases, age and their relative risks of accident involvement: results from a meta-analysis. TOI Report 690, 2003.
5. Charlton J, Koppel S, O'Hare M *et al.* Influence of chronic illness on crash involvement of motor vehicle drivers. Victoria, Australia: Monash University Accident Research Centre, 2004. Report No. 213.
6. British Columbia Medical Association. 2010 BC Guide in Determining Fitness to Drive. Victoria, Canada: Ministry of Public Safety and Solicitor General Office of the Superintendent of Motor Vehicles, 2010.
9. Etgen T, Sander D, Chonchol M *et al.* Chronic kidney disease is associated with incident cognitive impairment in the elderly: the INVADE study. Nephrol Dial Transplant 2009; 24: 3144–50.
12. Cacciatore F, Abete P, Calabrese C *et al.* Osservatorio Geriatrico Campano Study Group. Congestive heart failure and cognitive impairment in an older population. J Am Geriatr Soc 1998; 46: 1343–8.

15. Saczynski JS, Jonsdottir MK, Garcia ME *et al*. Cognitive impairment: an increasingly important complication of type 2 diabetes. Am J Epidemiol 2008; 168: 1132–9.

17. Potter GG, Steffens DC. Contribution of depression to cognitive impairment and dementia in older adults. Neurologist 2007; 13: 105–17.

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

22. Fitten LJ, Perryman KM, Wilkinson CJ *et al*. Alzheimer and vascular dementias and driving: a prospective road and laboratory study. JAMA 1995; 273: 1360–5.

30. Friedland RP, Koss E, Kumar A *et al*. Motor vehicle crashes in dementia of the Alzheimer type. Ann Neurol 1988; 24: 782–6.

31. Gilley DW, Wilson RS, Bennett DA *et al*. Cessation of driving and unsafe motor vehicle operation by dementia patients. Arch Intern Med 1991; 151: 941–6.

32. Lucas-Blaustein MJ, Filipp L, Dungan C, Tune L. Driving in patients with dementia. J Am Geriatr Soc 1988; 36: 1087–91.

34. Trobe JD, Waller PF, Cook-Flannagan CA, Teshima SM, Bieliauskas LA. Crashes and violations among drivers with Alzheimer disease. Arch Neurol 1996; 53: 411–6.

35. Ismail Z, Rajji TK, Shulman KI. Brief cognitive screening instruments: an update. Int J Geriatr Psychiatry 2010; 25: 111–20.

36. Mathias JL, Lucas LK. Cognitive predictors of unsafe driving in older drivers: a meta-analysis. Int Psychogeriatr 2009; 21: 637–53.

37. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12: 189–98.

38. Reitan RM. Trail Making Test: Manual for Administration, Scoring and Interpretation. Indianapolis, IN: Indiana University Medical Center, 1958.

39. Iverson DJ, Gronseth GS, Reger MA *et al*. Practice parameter update: evaluation and management of driving risk in dementia: report of the quality standards subcommittee of the American Academy of Neurology. Neurology 2010; 74: 45–65.

42. Marottoli RA, Richardson ED, Stowe MH *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: 562–8.

44. Ball KK, Roenker DL, Wadley VG *et al*. Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? J Am Geriatr Soc 2006; 54: 77–84.

45. Bedard M, Weaver B, Darzins P, Porter MM. Predicting driving performance in older adults: we are not there yet! Traffic Inj Prev 2008; 9: 336–41.

46. Staplin L, Gish KW, Wagner EK. MaryPODS revisited: updated crash analysis and implications for screening program implementation. J Saf Res 2003; 34: 389–97.

47. Staplin L, Lococo K, Gish K, Decina L. Model Driver Screening and Evaluation Program: Final Technical Report: Volume II: Maryland Pilot Older Driver Study. Washington, DC: National Highway Traffic Safety Administration, 2003. Report No. DOT HS 809 583.

48. Stutts JC, Stewart JR, Martell C. Cognitive test performance and crash risk in an older driver population. Accid Anal Prev 1998; 30: 337–46.

49. Mazer BL, Korner-Bitensky NA, Sofer S. Predicting ability to drive after stroke. Arch Phys Med Rehabil 1998; 79: 743–50.

53. Strauss E, Sherman EMS, Spreen O. Trail making test. In: A Compendium of Neuropschological Tests. New York: Oxford University Press, 2006; 655–9.

57. Barrash J, Stillman A, Anderson SW, Uc EY, Dawson JD, Rizzo M. Prediction of driving ability with neuropsychological tests: demographic adjustments diminish accuracy. J Int Neuropsychol Soc 2010; 16: 679–86.

60. Fuller R. Towards a general theory of driving behavior. Accid Anal Prev 2005; 37: 461–72.