Diagnostic Accuracy of the Chinese Version of the Trail-Making Test for Screening Cognitive Impairment

Mingqing Wei, MD,*a Jing Shi, MD,*a Ting Li, MM,* Jingnian Ni, MD,* Xuekai Zhang, PhD,* Yumeng Li, MD student,* Shenghua Kang, MD student,* Fuyun Ma, MD student,* Hengge Xie, MD,† Bin Qin, MD,‡ Dongsheng Fan, MD,§ Liping Zhang, MD,¶ Yongyan Wang, MD,** and Jinzhou Tian, MD, PhD*

BACKGROUND/OBJECTIVES: The Trail-Making Test (TMT), which is commonly used to measure executive function, consists of two components (TMT-A and TMTB). There is a lack of normative TMT data for Chinese elderly adults. This study aimed to evaluate the validity of the TMT in screening for cognitive impairment.

DESIGN: 2,294 Chinese-speaking adults aged 50 to 85: 1,026 with normal cognition (NC), 462 with mild cognitive impairment (MCI), 108 with Alzheimer’s disease (AD), 113 with vascular mild cognitive impairment (VaMCI), 121 with vascular dementia (VaD), 282 with uncertain types of dementia, and 15 with mixed dementia. Receiver operating characteristic curve analysis was performed to test the ability of TMT scores to differentiate between NC and cognitive impairment.

RESULTS: Age, education, and sex were significantly associated with TMT completion time. The TMT-A exhibited sensitivity of 77.8% and specificity of 92.0% with cut-off value of 98.5 seconds for discriminating AD from NC. The TMT-B had sensitivity of 83.3% and specificity of 91.8% with a cut-off value of 188.5 seconds for discriminating AD from NC. The TMT-A had sensitivity of 85.7% and specificity of 81.6% for discriminating NC from VaD with a cut-off value of 77.5 seconds, and the TMT-s had sensitivity of 81.6% and specificity of 83.9% with a cut-off value of 147.5 seconds. The TMT had less sensitivity distinguishing MCI from NC.

CONCLUSION: The Chinese version of the TMT is reliable for detecting AD or VaD but poor at distinguishing MCI from NC. J Am Geriatr Soc 66:92–99, 2018.

Key words: Trail-Making Test; clinical norms; vascular dementia; Alzheimer’s disease

With a rapidly aging population, dementia represents a globally major public health problem. Alzheimer’s disease (AD) is the most common type of dementia. Executive function impairment is considered a predictor of progression from mild cognitive impairment (MCI) to clinical AD and is thought to be associated with prefrontal dysfunction.1–3 The Trail-Making Test (TMT) consists of two components (A and B) and is among the most widely used neuropsychological assessment instruments as an indicator of cognitive processing speed and executive functioning.4 The TMT-A is widely used as a measure of psychomotor speed and visual attention, and the TMT-B measures cognitive alternation and set-shifting abilities.5,6 In the original English version of the TMT,7 the TMT-A requires subjects to draw a line consecutively through 25 digits randomly distributed on a sheet of paper, in ascending order, as quickly as possible. TMT-B comprises consecutive numbers and letters, and the goal is to connect numbers and letters alternately; because most Chinese older people are culturally less familiar with Arabic numerals than Westerners, a Chinese version of the TMT was developed,8 but there is little normative neuropsychological information on the TMT in elderly Chinese populations, and the most-appropriate cut-off values are unclear.

This study aimed to determine whether demographic factors affect TMT scores and to establish the optimal cut-off values for discriminating individuals with cognitive impairment from those with normal cognition (NC) in a Chinese population.
METHODS

Participants
Chinese-speaking adults aged 50 to 85 with memory complaints were screened between January 2013 and August 2015 in the memory clinic of Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China. Every participant underwent a standard neuropsychological assessment. Subjects were allocated to groups according to the results of a mental state examination, neuropsychological assessment, laboratory results, and neuroimaging findings.

Subjects meeting criteria from the Mayo Clinic for healthy controls were allocated to the normal control (NC) group. The criteria were no active neurological or psychiatric diseases; no antipsychotic drug use; no medical disorder or a treated disorder that does not compromise cognitive function; Mini-Mental State Examination score (MMSE) greater than 26 (higher education), greater than 24 (middle school education), greater than 23 (primary school education), greater than 22 (illiterate) (based on a previous study in the Chinese population); Clinical Dementia Rating Scale score of 0; and no activity of daily living disability, the instrumental activities of daily living scale ADL below 16.

Individuals with MCI were required to meet the diagnostic criteria for MCI reported by the MCI Working Group of the European Consortium on Alzheimer’s Disease; individual or their caregiver has cognitive complaints; normal general cognitive function identified by a clinician; MMSE score of 24 or greater; cognitive disorder as evidenced by clinical evaluation or CDR score of 0.5; preservation of ability of daily living function (ADL <16); absence of dementia judged by an experienced dementia research clinician, including no impairment of cognitive function that would meet National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer’s Disease and Related Disorders Association workgroups criteria for AD; and a score of 4 or less on the Hachinski Ischemic Scale (HIS) or 12 or fewer of 17 items on the Hamilton Depression Rating Scale.

Participants were also required to have no or minimal medial lobe temporal atrophy (MTA; including hippocampal volume atrophy) on magnetic resonance imaging (MRI), evaluated using the medial temporal lobe atrophy score.

The diagnostic criteria for probable AD were based on the core clinical criteria of the National Institute on Aging—Alzheimer’s Association workgroups. The inclusion criteria were as follows: (1) gradual and progressive cognitive impairment of cognitive function severe enough to interfere with ADL function (ADL <16); cerebrovascular disease (HIS score > 4); and correlation between dementia and cerebrovascular disease: (a) dementia occurred within 3 months after a stroke or (b) evidence of sudden onset, stepwise progression, and focal cortical deficits on neuropsychological assessment.

Subjects with vascular mild cognitive impairment (VaMCI) met the guidelines for diagnosis of vascular MCI: impairment in one or more cognitive domains verified by neuropsychological assessment; normal general cognitive function identified by a clinician, based on a systematic interview of the participant (MMSE score ≥24); cerebrovascular disease confirmed by MRI; cognitive impairment associated with the cerebrovascular events within 3 months or a HIS score of 7 or greater; preservation of ADL function (ADL score < 16); absence of dementia; and exclusion of other reasons for cognitive disorder. Exclusion criteria were meeting criteria for dementia, having depression or other psychosis, and having no evidence of cerebrovascular diseases on computed tomography or MRI.

All inclusion criteria for MCI, AD, and VaMCI have been described previously. The diagnostic pathway is shown in Supplementary Figure S1.

Chinese Version of the TMT
The Chinese version of the TMT was developed previously. The TMT consists of two components. The TMT-A consists of 25 consecutive numbers, from 1 to 25. The TMT-B was modified as 25 numbers enclosed in 13 circles (from 1 to 13) and 12 squares (from 1 to 12), as part of adapting the test to Chinese populations. Instructions for performing and scoring the test were translated into Chinese and then back into English for verification. The Chinese version of TMT is shown in Supplementary Figure S2.

Procedures
The TMT is administered as follows.

TMT-A
Individuals are instructed to draw a line as rapidly as possible joining consecutive numbers (i.e., 1–2–3... 24–25) randomly arranged on an A4 page (21 × 29 cm). A maximum time of 150 seconds was allowed for completion.

TMT-B
Individuals are required to draw a line alternating between circles and squares while connecting the numbers in
| Group                          | Normal Control, n = 1,026 | Mild Cognitive Impairment, n = 462 | Vascular Cognitive Impairment, n = 113 | Alzheimer’s Disease, n = 108 | Vascular Dementia, n = 122 | Mixed Dementia, n = 15 | Uncategorized Dementia, n = 282 | F/Chi-Square | P-Value |
|-------------------------------|---------------------------|-----------------------------------|----------------------------------------|-----------------------------|---------------------------|------------------------|-------------------------------|--------------|---------|
| Age, mean ± SD                | 66.2 ± 8.6                | 68.0 ± 9.36                       | 67.1 ± 9.4                             | 71.5 ± 8.7                  | 70.8 ± 10.4               | 73.3 ± 6.7              | 73.6 ± 8.5                   | 34.5         | <.001   |
| Education, years mean ± SD    | 12.1 ± 3.6                | 11.1 ± 3.6                       | 11.2 ± 3.8                             | 11.7 ± 4.2                  | 10.6 ± 4.0                | 10.5 ± 3.9              | 10.1 ± 4.8                   | 11.2         | <.001   |
| Sex, n                        |                           |                                   |                                        |                             |                           |                        |                               |              |         |
| Male                          | 362                       | 181                               | 62                                      | 46                          | 69                        | 8                      | 132                           | 70.0         | <.001   |
| Female                        | 664                       | 281                               | 51                                      | 62                          | 53                        | 7                      | 150                           |              |         |
| Mini-Mental State Examination, mean ± SD | 28.2 ± 1.6               | 26.0 ± 2.4                        | 25.8 ± 2.2                             | 15.0 ± 6.6                  | 19.2 ± 6.9                | 13.5 ± 6.4              | 17.0 ± 5.8                   | 542.4        | <.001   |
| Clock Drawing Test, mean ± SD | 3.9 ± 0.4                 | 3.5 ± 0.7                         | 3.5 ± 0.9                              | 1.8 ± 1.4                   | 2.3 ± 1.3                 | 1.7 ± 1.6              | 2.0 ± 1.3                    | 274.7        | <.001   |
| Delayed story recall, mean ± SD | 25.6 ± 8.7              | 8.4 ± 7.5                         | 9.3 ± 7.8                              | 0.6 ± 2.2                   | 6.0 ± 9.2                 | 1.3 ± 4.8              | 1.3 ± 4.0                    | 452.9        | <.001   |
| Activity of daily living score, mean ± SD | 14.1 ± 1.3              | 14.7 ± 2.5                        | 14.7 ± 1.6                             | 27.0 ± 9.7                  | 26.1 ± 10.2               | 30.2 ± 10.0             | 25.5 ± 8.8                   | 275.1        | <.001   |
| Trail Making Test, mean ± SD  |                           |                                   |                                        |                             |                           |                        |                               |              |         |
| Part A                        | 61.0 ± 26.5               | 79.0 ± 33.3                       | 89.1 ± 35.0                            | 126.5 ± 35.6               | 119.2 ± 35.9             | 134.8 ± 27.5             | 125.6 ± 32.0                  | 239.0        | <.001   |
| Part B                        | 106.6 ± 57.6              | 163.4 ± 85.1                      | 166.7 ± 82.0                           | 260.0 ± 76.6               | 234.2 ± 82.5             | 262.5 ± 69.0             | 250.5 ± 76.5                  | 181.8        | <.001   |

SD = standard deviation.
ascending order (i.e. ①→②→③→, etc.) as rapidly as possible on a page of the same size. A maximum time of 300 seconds was allowed for completion.

Before each part of the test, a practice trial with eight items was administered to ensure that individuals understood each task. During the tests, the examiner corrected errors immediately and instructed individuals to return to the last correct target. The time taken to complete each part of the task (in seconds) was recorded as the score. Errors defined as any incorrect line that reached the target. The TMT score had no effect on diagnostic decisions about the participants. The Dongzhimen Hospital, Beijing University of Chinese Medicine Institutional Ethics Committee approved the protocol. The study was undertaken in accordance with the principles of the Declaration of Helsinki. Participants and responsible caregivers provided written informed consent.

Statistical Methods

SPSS version 20.0 for Windows (IBM Corp., Armonk, NY) was used for data analyses. Analysis of variance was performed to compare mean ages, education, and neuropsychological scale scores between the groups. Sex distributions were compared using chi-square tests. Multiple regression analysis was used to examine the influence of age and education on TMT completion time. Receiver operating characteristic (ROC) curve analysis was performed to calculate the optimal sensitivity and specificity using different cut-off values of the TMT; P < .05 was considered statistically significant.

RESULTS

Demographic and Neuropsychological Variables

Two thousand four hundred forty-nine individuals were enrolled, and 155 were excluded (130 unable to complete the neuropsychological assessment, 25 declined to participate), leaving 2,294 subjects for the final analysis. All subjects were tested on TMT-A and TMT-B; all completed the TMT-A, and 2,086 completed the TMT-B (208 did not understand the test instructions); and 1,026 were diagnosed with NC, 462 with MCI, 108 with AD, 113 with depression, 15 with VaMCI, 122 with VaD, 166 with mixed dementia, and 282 uncategorized dementia because of lack of MRI. The study subject flow chart is shown in Supplementary Figure S3.

The demographic characteristics of the participant groups are shown in Table 1. There were significant differences in age and education between groups. The AD and VaD groups were significantly older than the NC (all P < .001), MCI (P < .001, P = .002) and VaMCI (all P < .001) groups. The NC group also had significantly more education than the other groups (all P < .001). There was a significant difference in the balance between the sexes between the two groups (P < .001).

Completion times for either TMT in the AD and VaD groups were significantly different from those of the NC (all P < .001), MCI (all P < .001), and VaMCI (all P < .001) groups.

Table 2. Normative Data for Trail-Making Test Part A (TMT-A) and Part B (TMT-B) Completion Times Stratified According to Age and Education

| Age | TMT-A | TMT-B |
|-----|-------|-------|
| 0-6 | Mean | SD  | IQR | Mean | SD  | IQR |
| 7-9 | n | | | | | |
| 50-64 | 126 | 51.3±21.5 | 27.0±9.0 | 61.6±22.0 | 32.8±8.4 | 33.3±10.1 |
| | 319 | 94.4±48.7 | 49.0±19.2 | 113.2±47.5 | 51.5±24.7 | 51.9±24.7 |
| ≥10 | 471 | 82.7±36.4 | 40.0±13.1 | 109.8±48.3 | 31.7 | 30.0±21.0 |

SD = standard deviation; IQR = interquartile range.
Effects of Demographic Factors on TMT Performance

Information on age, education, and sex was entered into a multiple linear regression analysis, with TMT-A and TMT-B as the dependent variables. Both models were statistically significant overall (\(F = 244.36, \ P < .001\); \(F = 215.86, \ P < .001\)), accounting for 24.2% and 23.7% of the variance in TMT-A and TMT-B completion times, respectively. Age (correlation coefficient (\(r\)) = 0.40, \(P < .001\)), education (\(r = -0.24, \ P < .001\)) and sex (\(r = 0.07, \ P < .001\)) were significantly correlated with TMT-A completion time, and these variables also affected TMT-B completion time (age: \(r = 0.38, \ P < .001\); education: \(r = -0.26, \ P < .001\); sex: \(r = 0.08, \ P < .001\)).

Normative Data for the Chinese Version of the TMT

Considering the significant effects of age and education on TMT, we divided the NC group into three age groups (50–64, 65–74, 75–85). Each age group was then divided into three subgroups based on years of education. These data are presented in Table 2.

TMT Completion Time for Discriminating of MCI from NC

ROC analysis was performed to provide cut-off TMT values for distinguishing subjects with MCI from NC (Figure 1A). The area under the ROC curve (AUC) was 0.66 (\(P < .001\)) (95% confidence interval (CI) = 0.63–0.69) for TMT-A and 0.70 (95% CI = 0.67–0.73) for TMT-B. When the cutoff score was 72.5 seconds for the TMT-A, the sensitivity and specificity was 48.4% and 78.4% for distinguishing MCI from NC. The sensitivity (51.8%) and specificity (80.2%) was obtained with the cutoff score 135.5 seconds for screening MCI from NC.

TMT Completion Time for Discriminating AD from NC

ROC analysis was performed to calculate the cutoff score and diagnostic value of TMT for discriminating AD from NC (Figure 1B). The AUC was 0.89 (\(P < .001\)) (95% CI = 0.85–0.93) for TMT-A and 0.91 (95% CI = 0.87–0.94) for TMT-B. When the cutoff values were 98.5 seconds for TMT-A, the sensitivity (77.8%) and specificity (92.0%) was obtained for distinguishing AD from NC. And the TMT-B showed higher sensitivity (83.3%) and specificity (91.8%) with cutoff 188.5 seconds for screening AD from NC.

The prevalence of MCI was 20.1% and of AD was 4.7%. The positive predictive values (PPVs) and negative predictive values (NPVs) of TMT-A were 0.20 and 0.80 for the detection of MCI and 0.32 and 0.99 for detecting AD. For TMT-B, the PPVs and NPVs were 0.40 and 0.87 for detecting MCI and 0.33 and 0.99 for detecting AD.

Considering the significant effects of age and education on the TMT, the cutoff score and diagnostic value of TMT for screening AD from NC was calculated according to different age and education group (Table 3).

TMT Completion Times for Discriminating VaMCI and VaD from NC

The sensitivity and specificity of TMT-A and TMT-B were determined using ROC analysis for discrimination of NC from VaMCI and VaD. The results revealed that the ability of TMT scores to discriminate between NC and VaD (AUC = 0.88 for TMT-A, AUC = 0.87 for TMT-B) was more robust than between NC and VaMCI (AUC = 0.74 for TMT-A, AUC = 0.72 for TMT-B).

When the cutoff score was 63.5 seconds for TMT-A, the sensitivity and specificity was 70.5% and 67.7% to
discriminating VaMCI from NC. For TMT-B, when the cutoff score was 126.5 seconds, the sensitivity (62.9%) and specificity (75.9%) was obtained for discriminating VaMCI from NC. Moreover, the TMT-A showed greater sensitivity (85.7%) and specificity (81.6%) in differentiating NC from VaD with the cutoff score 77.5 seconds, and the TMT-B appeared to higher sensitivity (81.6%) and specificity (83.9%) in differentiating between NC and VaD when the cutoff values were 147.5 seconds.

Partial Correlation Between MMSE and TMT

Because age and education affected TMT completion time, we calculated the correlation between MMSE scores, MMSE sub domain scores, MMSE include sub domain including orientation, instant recall, attention and calculation, language, delayed recall and TMT completion time while controlling for age and education with the partial correlation analysis. Completion times for TMT-A ($r = -0.6$, $P < .001$) and TMT-B ($r = -0.6$, $P < .001$) were significantly correlated with MMSE scores. All of the sub domains were significantly correlated with the TMT-A and TMT-B, and the correlation coefficient for orientation was greater than 0.5, the other domains were relative low (below than 0.5).

DISCUSSION

A previous study reported that the Chinese version of the TMT was applicable for the Chinese population and could be used to identify individuals with mild AD, but because of the limited sample size, a stratification analysis of educational level was not conducted. Thus, the optimal cut-off values for screening MCI and AD using the TMT remained unclear. The current study was conducted to examine the influence of demographic factors on TMT performance. The multiple linear regression analysis in the study demonstrated that age, education, and sex influenced both parts of the TMT. Age was positively correlated with TMT scores, whereas education level was negatively correlated. Sex also had a faintly positive correlation with TMT performance. The results showed that older women and individuals with less education took more time to complete the test. This finding is consistent with other studies conducted in populations with different languages and cultural backgrounds. In the current study, age had a stronger effect than education, consistent with previous reports.

Because age and education are related to TMT performance, this study examined normative data of TMT stratified according to age and education in the NC group. A previous study conducted in the Czech Republic reported norms for TMT-A performance, with norm values of 35.4 ± 13.2 for individuals aged 55 to 64, 40.3 ± 12.4 for those aged 65 to 74, and 46.8 ± 20.9 for those aged 75 to 84. Completion times for the TMT-A and TMT-B in our sample were markedly longer than those reported with the Czech version. This difference may be attributable to cultural differences between the two countries; there are two digital systems in China: Arabic numerals and traditional Chinese numerals, and most Chinese older people are less familiar with the Arabic numerals than...
The results are consistent with those of an earlier study. MCI and VaMCI from NC were low, indicating that TMT specificity of TMT for detecting AD was more than 90% in the current study. Nonetheless, the accuracy in discriminating MCI and VaMCI from NC were low, indicating that TMT may not be suitable for screening for MCI and VaMCI. The results are consistent with those of an earlier study conducted in a Chinese population.

In the current study, TMT-B performance was poorer than TMT-A performance; 208 subjects failed to complete TMT-B because of inability to understand the instructions, and many participants with AD and VaD took up to 300 seconds to complete TMT-B. This may have been because TMT-B has a more complex physical layout and greater task demands and is more difficult to process and complete than TMT-A.

In the Chinese version, we only replaced the letters with numbers. Unlike with the English version, the presence of the same number in the Chinese version of the TMT-B may be a cue, even though the surrounding shapes changes. In addition, conducting this study in only one center may have introduced selection bias. Hence, a larger-scale, multicenter study should be conducted to evaluate the utility of TMT in Chinese elderly adults.

CONCLUSION

The TMT is considered a useful screening tool for detecting AD and VaD with high accuracy in the Chinese population. It does not appear to be suitable for distinguishing MCI and VaMCI from normal cognition. It is recommended that cultural variables and demographic factors be taken into consideration when administering and interpreting the results of the TMT.

ACKNOWLEDGMENTS

Conflict of Interest: There are no competing interests.

Author Contributions: JT and JS designed the study and planned analysis of the data. MW wrote the manuscript. TL, JN, XZ, SK, FM, RZ, ZC, DW, XL, DF, H X performed the neuropsychological assessment. LZ administered MRI. JT and JS were principal investigators for this study and finalized the manuscript. BQ revised the manuscript. YW reviewed the design of the trial protocol. All authors reviewed and approved the final version.

Sponsor’s Role: Data collection was supported by Grant B08006 from the “111 Project,” Grants 81473518 and 81573824 from the National Natural Science Foundation of China, Grants SF2014–1-4191 and SF2016–4-4193 from the Capital Health Research and Development of Special, and Grants Z141107002515019 and Z15110003815021 from the Beijing Municipal Science and Technology Commission Foundation, China.

REFERENCES

1. Boyle PA, Malloy PF, Salloway S et al. Executive dysfunction and apathy predict functional impairment in Alzheimer disease. Am J Geriatr Psychiatry 2003;11:214–221.
2. Chen P, Ratchiff G, Belle SH et al. Cognitive tests that best discriminate between presymptomatic AD and those who remain non demented. Neurology 2000;55:1847–1853.
3. Chen ST, Sutzer DL, Hinkin CH et al. Executive dysfunction in Alzheimer’s disease: Association with neuropsychiatric symptoms and functional impairment. J Neuropsychiatry Clin Neurosci 1999;10:426–432.
4. Sánchez-Cubillo L, Pérez-Ja, Adrover-Roig D et al. Construct validity of the trail making test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J Int Neuropsychol Soc 2009;15:438–450.
5. Hashimoto R, Meguro K, Lee E et al. Effect of age and education on the Trail-Making Test and determination of normative data for Japanese elderly people: The Tajiri Project. Psychiatry Clin Neurosci 2006;60:422–428.
6. Stuss DT, Bisschop SM, Alexander MP et al. The trail making test: A study in focal lesion patients. Psychol Assess 2001;13:230–239.
7. Partington JE, Leiter RG. Partington’ pathway test. Psychological Service Bulletin 1949;1:9–20.
8. Lu JC, Guo QH, Hong Z et al. Trail making test used by Chinese elderly patients with mild cognitive impairment and mild Alzheimer’ dementia. Chinese J Clin Psycho 2006;14:118–120 Chinese.
9. Ivnik R, Malec JF, Smith GE et al. Mayo’s older Americans normative studies: WAIS-R norms for ages 56 to 97. Clin Neuropsychol 2007;6:1–30.
10. Folstein MF, Folstein SE, McHugh PR. ‘Mimic-mental state’. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1974;12:189–198.
11. Tangalos EG, Smith GE, Ivnik RJ et al. The Mini-Mental State Examination in general medical practice: Clinical utility and acceptance. Mayo Clin Proc 1996;71:829–837.
12. Hughes CP, Berg I, Danziger WL et al. A new clinical scale for the staging of dementia. Br J Psychiatry 1982;140:366–572.
13. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9:179–186.
14. Portet F, Osusij PJ, Visser PJ et al. Mild cognitive impairment (MCI) in medical practice: A critical review of the concept and new diagnostic procedure. Report of the MCI Working Group of the European Consortium on Alzheimer’s Disease. J Neurol Neurosurg Psychiatry 2006;77:714–718.
15. Tian J, Liang X, Shi J et al. Diagnosis and Treatment of Alzheimer’s Disease. Beijing, China: People’s Health Press, 2010.
16. McKhann GM, Knopman DS, Chertkow H et al. The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimers Dement 2011;7:263–269.
17. Hachinski VC, Liff LD, Zilhka E et al. Cerebral blood flow in dementia. Arch Neurol 1975;35:632.
18. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62.
19. Ni J, Shi J, Wei M et al. Screening mild cognitive impairment by delayed story recall and instrumental activities of daily living. Int J Geriatr Psychiatry 2015;30:888–890.
20. Román GC, Tatami T, Erkinjuntti T et al. Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 1999;53:250–260.
21. Gorelick PB, Scuteri A, Black SE et al. Vascular contributions to cognitive impairment and dementia: A statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011;42:2762–2773.
22. Shi J, Tian J, Wei M et al. The utility of the Hopkins Verbal Learning Test (Chinese version) for screening dementia and mild cognitive impairment in a Chinese population. BMC Neurol 2012;12:136.
23. Shi J, Wei M, Tian J et al. The Chinese version of story recall: A useful screening tool for mild cognitive impairment and Alzheimer’s disease in the elderly. BMC Psychiatry 2014;14:71.
24. Perić J, Rios-Lago M, Rodríguez-Sánchez JM et al. Trail Making Test in traumatic brain injury, schizophrenia, and normal ageing: Sample comparisons and normative data. Arch Clin Neuropsychol 2007;22:433–447.
25. Woods DL, Wyma JM, Herron TJ et al. The effects of aging, malingering, and traumatic brain injury on computerized Trail-Making Test performance. PLoS ONE 2015;10:e0124345.

26. Tombaugh TN. Trail Making Test A and B: Normative data stratified by age and education. Arch Clin Neuropsychol 2004;19:203–214.

27. Bezdicke O, Motak L, Axelrod BN et al. Czech version of the Trail Making Test: Normative data and clinical utility. Arch Clin Neuropsychol 2012;27:906–914.

28. Ashendorf L, Jefferson AL, O’Connor MK et al. Trail Making Test errors in normal aging, mild cognitive impairment, and dementia. Arch Clin Neuropsychol 2008;23:129–137.

29. Seo EH, Lee DY, Kim KW et al. A normative study of the Trail Making Test in Korean elders. Int J Geriatr Psychiatry 2006;21:844–852.

**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

- **Figure S1.** The Chinese version of Trail making test.
- **Figure S2.** Diagnostic algorithm for AD, MCI, VaD, VaMCI and normal cognition.
- **Figure S3.** Standard study flow chart. Notes: MCI = mild cognitive impairment; AD = Alzheimer’s disease; VaMCI = vascular cognitive impairment; VaD = vascular dementia.

Please note: Wiley-Blackwell is not responsible for the content, accuracy, errors, or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.