Executive dysfunction (ED) is a common neuropsychological feature associated with subcortical ischemic vascular disease. It is due to the disruption of frontal-subcortical circuits linking the dorsal-lateral prefrontal cortex. Executive functions encompass a host of higher cognitive processes including abstraction, initiation, planning, sustained attention, perseveration, set shifting and monitoring of progress. Among stroke survivors, ED was found to be associated with poor performance in both basic and complex activities of daily living.

Evaluation of the presence and severity of ED provides clinicians important information for treatment and rehabilitation planning.

Although tests for executive functions such as the Wisconsin card sorting test and trail making test are available, they are too extensive to be used as screening instruments. The traditional cognitive screening test, the Mini Mental State Examination is biased towards detection of posterior cortical deficits such as memory and language disturbance, and is not sensitive in evaluating ED. To date, a valid and reliable screening test designed specifically for detecting ED is not yet available for our local elderly population.

The executive clock drawing task (CLOX) is a poor screening test for executive dysfunction in Chinese elderly patients with small subcortical ischemic vascular disease

Summary

Executive dysfunction (ED) is a prominent feature of subcortical ischemic vascular disease. A screening test for ED is lacking among Chinese. The objective of the study is to investigate the validity and reliability of a Chinese version of the Executive clock drawing task (CLOX) in screening ED among Chinese elderly patients with small subcortical infarct (SSI). The Chinese version of CLOX correlated with MMSE, CDRS I/P, and WCST perseverative errors. However, multivariate regression analysis showed that only education (R² change = 0.22, p < 0.001) and MMSE (R² change = 0.35, p < 0.001), but none of the standard executive function tests, significantly accounted for the variance in the CLOX. Test–retest (r = 0.84) and inter-rater reliability (r = 0.84) were high for the CLOX.

Conclusions. Although the CLOX is reliable, it is not valid in detecting ED in Chinese elderly patients with SSI.

Keywords: small subcortical infarcts, executive dysfunction, CLOX, chinese
center (n = 30), relatives of hospital staff (n = 6), spouse of patients (n = 3) and hospital staff (n = 2). Criteria for inclusion as controls were normal vision and hearing, absence of any past or present neurological, or psychiatric illness, and no cognitive symptoms. All controls had a clinical dementia rating (CDR) of 0 and were independent in activities of daily living.

Thirty patients were recruited randomly from our stroke clinic. Stroke patients with SSI seen in cerebral MRI or CT were eligible to participate in the study. Small subcortical infarct was defined as a hyperintense (T2 weighted MRI) or hypodense (CT) lesion of size between 0.2 and 2 cm in all dimensions that was located in the cerebral white and deep gray matter, and the white matter of the cerebellum. The sites of symptomatic SSI were recorded and were classified into cerebral white matter, striatocapsular, thalamus, and cerebellum based on neuroimaging and clinical presentation. Patients with cortical, large subcortical, or brainstem infarct, and past or present features of intracerebral hemorrhage, non-ischemic lesions, e.g. tumor or demyelination, were excluded from the study. Inclusion of the majority of the patients was based on MRI (n = 28) and inclusion of two patients was based on CT. The same radiologist (WWML) read all the neuroimaging. Furthermore, patients with severe motor impairment interfering with participation in cognitive tests, stroke event occurred within 3 months of the study, cognitive impairment attributed to other illnesses, e.g. vitamin B12 deficiency or chronic alcoholism, and DSM-IV criteria of major depression or schizophrenia were also excluded from the study.

A semi-structured clinical interview was then performed on all recruited patients who had SSI and on their close informants to assess the presence of cognitive symptoms and the temporal relationship between cognitive symptoms and stroke. The CDR was also performed on all the patients during the interview. Special caution was made to grade functional impairment that was attributed to cognitive dysfunction rather than to motor weakness due to stroke or other medical illness. Patients who had acute deterioration of cognitive symptoms post-stroke and a CDR of ≥ 0.5 were labeled as subcortical ischemic cognitive impairment. The cognitive impairment of this group of patients was likely attributed to the SSI given the temporal relationship between cognitive symptoms and the stroke event. The severity of subcortical ischemic cognitive impairment was further graded as mild if the CDR = 0.5, and as severe if the CDR ≥ 1. Patients with no cognitive symptoms pre- or post-stroke were labeled as no cognitive impairment. Patients with progressive pre-stroke memory loss and who denied further stepwise deterioration of cognitive symptoms post-stroke were labeled as degenerative cognitive impairment. The same neurologist (VM) performed this interview. Stroke severity was graded according to the National Institute of Health Stroke Scale (NIHSS). All controls and patients gave their written consent to participate in this study and were administered the CLOX, Chinese MMSE (CMMSE), Chinese version of Mattis dementia rating scale (CDRS I/P), WCST, Barthel index (BI), and Lawton instrumental activities of daily living (IADL). Classification of CI and sites of SSI of patients are shown in Table 1.

### Statistics

Comparisons of demographic information were made using independent samples t test for continuous variables and χ² tests for categorical variables. Analysis of covariance (ANCOVA) was used to compare the performance of neuropsychological tests between healthy controls and patients with age and education entered as covariates. Pearson correlation coefficient was used to assess the correlations between the CLOX subscores and MMSE, CDRS I/P, WCST category completed and perseverative errors. Two separate multivariate regression analyses with forward stepwise method were carried out to assess the contributions of the MMSE, CDRS I/P, WCST category completed and perseverative errors to CLOX 1 and CLOX 2 with age and education entered in the first step. Test-retest and inter-rater reliability were assessed for each CLOX subscore using intraclass correlation coefficient. Three separate canonical discriminant function analyses were performed to assess the ability of CLOX 1, CLOX 2 and CLOX 1 + CLOX 2 to discriminate between healthy controls and patients with subcortical ischemic cognitive impairment. For the purpose of comparison, a fourth discriminant function was carried out using MMSE. Test retest and inter-rater reliability was assessed by the intraclass correlation between the scores of two administrations of the CLOX separated by a two weeks period. The same trained research assistant (AW) administered the CLOX, CDRS I/P, WCST, and MMSE. Another trained research assistant participated only by performing the CLOX for the evaluation of inter-rater reliability. Statistical significance was set at p ≤ 0.05. All statistical analyses were performed using SPSS version 11 (SPSS, Inc.).

### RESULTS

Healthy controls and patients did not differ in age, education, gender, and years of education. The two groups differed in BI (mean difference H–SICI = 1.96, p < 0.01) and IADL (mean difference H–SICI = −1.05, p < 0.001) (Table 2). Healthy controls performed better than SSI patients on all neuropsychological tests (p = 0.047 to p < 0.001) (Table 3). Three patients refused to perform the CLOX despite having adequate power of their dominant hands.
Discriminant validity

Discriminant analyses for CLOX 1, CLOX 2 and CLOX 1+CLOX 2 were all significant. Both CLOX 1 and CLOX 2 correctly classified 73.8% of subjects (CLOX 1: Wilk’s Lambda = 0.697, p < 0.001; CLOX 2: Wilk’s Lambda = 0.711, p < 0.001). CLOX 1 has a higher sensitivity (CLOX 1 = 65%; CLOX 2 = 40%) and a lower specificity (CLOX 1 = 78%; CLOX 2 = 90%) than CLOX 2. Among the CLOX scales, CLOX 1+CLOX 2 yielded the highest discriminant ability (Wilks’ Lambda = 0.670, p < 0.001). Interestingly, the discriminant ability of MMSE (Wilks’ Lambda = 0.664, p < 0.001) was almost identical to that of CLOX 1+CLOX 2 and both functions correctly classified 77% of subjects with a sensitivity of 65% and specificity of 82.9%. Details of the discriminant analyses are shown on Table 7.

Concurrent validity

CLOX 1 correlated significantly with MMSE (r = 0.74, p < 0.01), CDRS I/P (r = 0.54, p < 0.01), and WCST perseverative errors (r = -0.31, p < 0.05). CLOX 2 correlated significantly with MMSE (r = 0.83, p < 0.01), CDRS I/P (r = 0.65, p < 0.001), WCST category completed (r = 0.29, p < 0.05), and WCST perseverative errors (r = -0.34, p < 0.01). Age and education also correlated with both CLOX subscales (Table 4).

Multivariate forward stepwise regression analysis showed that the variance in CLOX 1 was explained by MMSE (R² change = 0.351, p < 0.001), age (p = NS), and education (R² change = 0.217, p < 0.001), but not by any of the executive measures. The variance in CLOX 2 was explained by MMSE (R² change = 0.45, p < 0.001), WCST perseverative errors (R² change = 0.021, p = 0.02), age (R² change = 0.09, p = 0.007), and education (R² change = 0.170, p < 0.001) (Table 5). Test-retest reliability was high for both CLOX 1 (r = 0.84, p < 0.001) and CLOX 2 (r = 0.95, p < 0.001). Inter-rater reliability was also high for CLOX 1 (r = 0.84, p < 0.001) and CLOX 2 (r = 0.90, p < 0.001) (Table 6).

DISCUSSION

The CLOX was designed to evaluate ED. Both CLOX 1 and 2 were shown to correlate with executive tests among healthy subjects and patients with AD. The CLOX 1 was shown to be able to detect even mild ED among relatively healthy elderly controls. Our original proposal was that the CLOX might provide a practical means to screen for ED among patients with subcortical pathology, such as subcortical stroke. However, our present study only showed a modest correlation between the CLOX and the formal executive test measures among our controls and patients with SSI and these formal executive measures did not significantly account for the variance of CLOX 1 in the stepwise multiple regression analysis. The correlation of CLOX 1 with executive measures was even lower when compared with those of CLOX 2. We noted that MMSE strongly correlated with both CLOX 1 and 2 and it had the greatest contribution to the variance of the CLOX scores. The ability to discriminate between controls and patients with cognitive impairment associated with SSI via using CLOX 1, CLOX 2 or CLOX 1+CLOX 2 was poorer or similar to those of MMSE.

Our findings suggest that the validity of CLOX in evaluating ED in our patient sample was poor. A reason why the CLOX correlated strongly with executive measure in the original study and not in our present study may be due differences in the executive measures that were used to validate the CLOX in the two studies. In the original study the CLOX was validated against the Executive Interview (EXIT25).20 The EXIT25 is another bedside executive screening test rather than a formal executive test. It was found to correlate moderately with traditional executive measures such as the WCST (r = 0.52) and Trail Making Test part B (r = 0.64) in its original validation study. Interestingly, the correlation of EXIT25 with MMSE (r = −0.85) was noted to be even stronger than the correlations with the executive measures. Some researchers suggested that EXIT25 was also sensitive to functions that are not executive.21 On the contrary, because our present study used formal executive measures, the WCST and CDRS I/P as gold standard executive tests that provide a good coverage of executive functions (abstraction, planning, conceptualization, set-shifting, initiation, perseveration), we believe our findings suggest CLOX as poor for evaluating ED. In fact, the strong correlation between CLOX and MMSE and the similar discriminant ability of both tests suggested that CLOX was at most sensitive in evaluating posterior cortical deficits or global cognition rather than ED.

Our findings also raise the question whether CDTs are well suited for evaluating ED. We agree with Royall and colleagues8,12 that clock drawing requires an orchestration of executive functions. However, CDTs are probably not a pure executive task because successful completion of CDT is also heavily dependent

Table 3  Comparisons of neuropsychological tests between healthy controls and patients

<table>
<thead>
<tr>
<th>Neuropsychological tests</th>
<th>Healthy controls</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>27.7 ± 2.1</td>
<td>23.1 ± 5.2</td>
</tr>
<tr>
<td>CDRS I/P</td>
<td>34.0 ± 3.0</td>
<td>24.7 ± 7.4</td>
</tr>
<tr>
<td>WCST # category</td>
<td>2.1 ± 1.8</td>
<td>1.2 ± 1.3</td>
</tr>
<tr>
<td>WCST perseverative errors</td>
<td>36.8 ± 19.4</td>
<td>53.7 ± 21.0</td>
</tr>
<tr>
<td>CLOX 1</td>
<td>10.4 ± 3.4</td>
<td>7.0 ± 4.0</td>
</tr>
<tr>
<td>CLOX 2</td>
<td>13.2 ± 2.1</td>
<td>10.8 ± 3.6</td>
</tr>
</tbody>
</table>

Results compared using ANCOVA controlled for age and education and shown in mean ± SD.

Table 4  Correlations between CLOX subscores and executive tests

<table>
<thead>
<tr>
<th>R</th>
<th>CLOX 1</th>
<th>CLOX 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.30†</td>
<td>0.41**</td>
</tr>
<tr>
<td>Education</td>
<td>0.47**</td>
<td>0.41**</td>
</tr>
<tr>
<td>MMSE</td>
<td>0.74**</td>
<td>0.83**</td>
</tr>
<tr>
<td>CDRS I/P</td>
<td>0.54**</td>
<td>0.65**</td>
</tr>
<tr>
<td>WCST category completed</td>
<td>0.16</td>
<td>0.29*</td>
</tr>
<tr>
<td>WCST perseverative errors</td>
<td>-0.31†</td>
<td>-0.34**</td>
</tr>
</tbody>
</table>

† p < 0.05.  ‡ p < 0.01.

Table 5  Multivariate stepwise regression analysis

<table>
<thead>
<tr>
<th>R² change</th>
<th>CLOX 1</th>
<th>CLOX 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>NS</td>
<td>0.09**</td>
</tr>
<tr>
<td>Education</td>
<td>0.22**</td>
<td>0.17**</td>
</tr>
<tr>
<td>MMSE</td>
<td>0.35**</td>
<td>0.45**</td>
</tr>
<tr>
<td>CDRS I/P</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>WCST # of category completed</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>WCST # of perseverative errors</td>
<td>NS</td>
<td>0.02†</td>
</tr>
</tbody>
</table>

† p < 0.05.  ‡ p < 0.01.
on non-executive functions, as evident from the fact that most CDT studies showed high correlations with MMSE.22 Because MMSE is sensitive to impairment in memory and cortical functions over subcortical functions, it is possible that, in practice, memory and cortical functions such as constructional praxis and linguistic competence play a more determinant role on CDT performance. Clinically, although it is plausible to infer the presence of ED in patients with normal MMSE and impaired CDT performance, such a distinction is less clear in patients with impaired performance on both tests because in the latter case, it would be difficult to attribute the failure of CDT to executive or non-executive dysfunctions.

Functional imaging studies have shown mixed findings regarding the areas of brain activated during performance on CDTs. In a single photon emission computed tomography (SPECT) study,23 the left posterior temporal, as opposed to the prefrontal regional cerebral blood flow predicted subjects’ performance on a CDT similar to the CLOX, suggesting the determinant function of semantic memory.24 In a functional magnetic resonance imaging (fMRI) study,24 however, the CDT performance was related strongly to the activation of the parieto-frontal areas, suggesting the involvement of executive functions. Most likely, CDT elicits many cognitive processes sub-served not by a singly confined, but multiple regions of the brain activated in parallel. Further studies are needed to confirm our speculation.

Because performance on CDT depends heavily on education,25 another possible reason why CLOX correlated only fairly with formal executive measures might be due to the possibility that CLOX is too difficult for Chinese subjects who are less educated or illiterate. In our present cohort, almost one quarter of subjects had no formal education and the mean years of education of subjects had no formal education and the mean years of education were much lower than the subjects in the original study (5 years vs. >10 years). Three subjects who had low education (education = 0, 1, and 3 years) even refused the CLOX because this group of patients.

REFERENCES
13. Royall DR, Espino DV, Polk M et al. Validation of a Spanish translation of
CLOX for use in Hispanic samples: the Hispanic EPESE study. Int J Geriatr


16. Chiu HFK, Kee HC, Chung WS, Kwong PK. Reliability and validity of the
Cantonese version of Mini-mental state examination – a preliminary study. J

17. Chan AS, Chot MK, Salmon DP. The effects of age, education, and gender on
the Mattis dementia rating scale performance of elderly Chinese and American

18. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Maryland

19. Lawton MP, Brody EM. Assessment of older people: self-maintaining and

20. Royall DR, Maharin RK, Gray KF. Bedside assessment of executive cognitive


22. Shulman KI. Clock-drawing: is it the ideal cognitive screening test. Int J Geriatr

23. Ueda H, Kitabayashi Y, Narumoto J et al. Relationship between clock drawing
test performance and regional cerebral blood flow in Alzheimer's disease: a
single photon emission computed tomography study. Psychiatry Clin Neurosci


25. Ainslie NK, Murden RA. Effect of education on the clock-drawing dementia

26. Lam LCW, Chiu HFK, Ng KO et al. Clock-face drawing, reading the setting