



OPEN

Association between cognitive quotient test score and hippocampal volume: a novel, rapid application-based screening tool

Wataru Kasai¹, Tadahiro Goto^{2✉}, Yuki Aoyama¹ & Kenji Sato³

We aimed to develop a brief, preclinical test to screen the reduced hippocampal volume that is a marker of early dementia [Cognitive Quotient (CQ) test]. We performed an observational study of adult subjects who underwent brain MRI in seven institutions from February 2018 to May 2019. The CQ test consists of five components: (1) digits forward, (2) digits backward, (3) Stroop test, (4) simple calculation, and (5) mental rotation. The primary outcome measure was hippocampal volume. We separated the data into derivation (n = 322) and validation cohorts (n = 96). In the derivation cohort, we built two scoring systems using the results of CQ test (model 1 and 2). In the validation cohort, we validated the correlation of the scoring systems with hippocampal volume. In the derivation cohort, there was a moderate correlation between the scoring systems and hippocampal volume [e.g., correlation coefficient = 0.62 in model 1 (95% CI 0.44–0.75)]. Likewise, in the validation cohort, there was a moderate correlation between the scoring systems and hippocampal volume [e.g., correlation coefficient = 0.54 in model 2 (95% CI 0.38–0.67)]. In this analysis of 418 participants, the score of newly developed CQ test was correlated with hippocampal volume.

Dementia is a progressive disorder characterized by a cluster of symptoms and signs, including psychological, psychiatric, cognitive, activity of daily living impairments. Approximately 12 million people worldwide have dementia, and the number will increase to 25 million by 2040¹. Since the global healthcare burden of dementia is larger than that of stroke, heart disease, and cancer, the benefits of early recognition and interventions are remarkable².

With the increasing use of magnetic resonance imaging (MRI), hippocampal volume has emerged as an attractive marker of early dementia, which was not identified in traditional screening tools, which is often examined as paper-based, due to the ceiling effect^{3–7}. In the *National Institute on Aging–Alzheimer’s Association* reported that the pathophysiological changes in medial temporal lobe including hippocampus might be associated with probable Alzheimer disease dementia⁸. However, although the early changes in the hippocampal volume has potential as a marker to detect subclinical dementia^{3,4,9,10}, the availability and financial costs of MRI may limit its use as a screening test in the healthy population without any apparent cognitive problems. Therefore, there is a need to develop a brief, preclinical test to screen the reduced hippocampal volume suggesting hippocampal atrophy in the healthy population—a marker of early dementia that is not identified by traditional diagnosis tools due to a ceiling effect in healthy subjects¹¹.

To address this concern, we have developed the Cognitive Quotient (CQ) test—a brief, application-based tool on digital devices to screen healthy subjects who may have a small hippocampal volume and require further investigations to detect early dementia. We aimed to examine the relationship between the score of the CQ test (CQ score) and MRI-based hippocampal volume.

¹Splink Inc. Japan, Tokyo, Japan. ²Graduate School of Medical Sciences, University of Fukui, Fukui, Japan. ³Department of Surgery, Sado General Hospital, Niigata, Japan. ✉email: tag695@mail.harvard.edu

	Overall	Derivation cohort	Validation cohort
Variables	n = 418	n = 322	n = 96
Age, median (IQR)	73 (59–80)	75 (66–81)	55 (48–70)
Women	223 (53%)	190 (59%)	33 (34%)
Height, cm, median (IQR)	158 (151–167)	150 (157–165)	157 (166–174)
Weight, kg, median (IQR)	56 (49–65)	55 (48–63)	63 (54–72)
Body mass index			
< 18.5	30 (7%)	24 (7%)	6 (6%)
18.5–24.9	305 (73%)	236 (73%)	69 (72%)
25.0–29.9	73 (17%)	55 (17%)	18 (19%)
≥ 30.0	10 (2%)	7 (2%)	3 (3%)
Education level, years^a			
6 or 9	38 (9%)	28 (9%)	10 (10%)
12	104 (25%)	73 (23%)	31 (32%)
14	26 (6%)	19 (6%)	7 (7%)
16 or(and) over	77 (18%)	36 (11%)	41 (43%)
N/A	173 (41%)	166 (52%)	7 (7%)
Smoking history^b			
Smoking index ≥ 1000	20 (5%)	16 (5%)	4 (4%)
Smoking index < 1000	117 (28%)	79 (25%)	38 (40%)
Never smoker	281 (67%)	227 (70%)	54 (56%)
Alcohol history			
Daily	118 (28%)	76 (24%)	42 (44%)
Weekly	73 (17%)	49 (15%)	24 (25%)
Monthly	34 (8%)	30 (9%)	4 (4%)
Not drinking	193 (46%)	167 (52%)	26 (27%)
Number of persons living with			
0	57 (14%)	43 (13%)	14 (15%)
1	141 (34%)	113 (35%)	28 (29%)
2	105 (25%)	81 (25%)	24 (25%)
3 or more	115 (28%)	85 (26%)	30 (31%)
Comorbidity			
Stroke including subclinical lacunar infarction	35 (8%)	35 (11%)	0 (0%)
Hypertension	52 (12%)	52 (16%)	0 (0%)
Diabetes	35 (8%)	35 (11%)	0 (0%)
Dementia or Alzheimer's disease	89 (21%)	89 (28%)	0 (0%)
Mild cognitive impairment	11 (3%)	11 (3%)	0 (0%)
Depression	4 (1%)	4 (1%)	0 (0%)
Missing	214 (28%)	118 (37%)	96 (100%)
Mini-mental status examination			
0–23	33 (8%)	33 (10%)	–
24–27	34 (8%)	34 (11%)	–
28–30	44 (11%)	44 (14%)	–

Table 1. Patient characteristics in derivation cohort and validation cohort. Data were expressed as n (%) unless otherwise indicated. Percentages may not equal 100 due to rounding. *IQR* interquartile range, *N/A* no answer. ^aEducation level was converted to “years of education” for each school category in questionnaire. ^bSmoking index (Brinkman index) was calculated as “smoking years-x-daily count”.

Results

Characteristics of the participants. From February 2018 to May 2019, a total of 449 participants completed the CQ test and underwent MRI. Among these participants, we excluded 10 participants with acute stroke and 21 participants whose hippocampal volumes could not be successfully abstracted from MRI artifacts. The remaining 418 participants were eligible for the analysis. The median age was 73 years, 53% of the participants were women, 73% had a normal body mass index, 67% had never smoked, and approximately 85% lived with ≥ 1 person (Table 1). Compared with the subjects in the derivation cohort, those in the validation cohort were more likely to be younger and male.

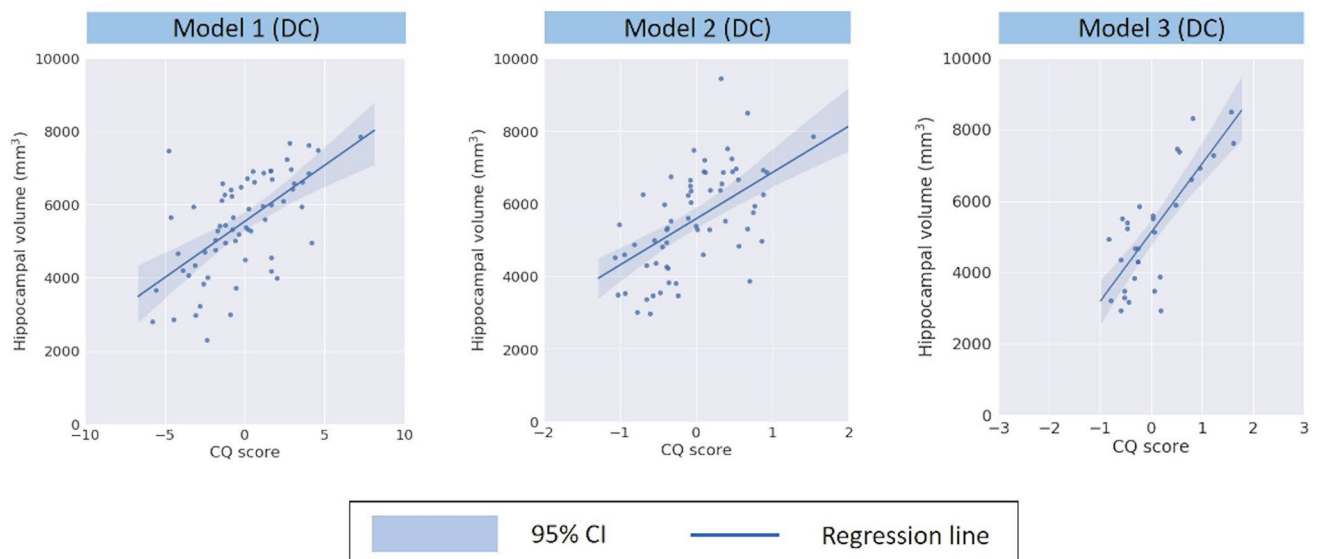


Figure 1. Correlation between Cognitive Quotient (CQ) score and hippocampal volume in the derivation cohort. In the derivation cohort, the model 1 has a moderate correlation between CQ score and hippocampal volume (mean r of the fivefold cross validation, 0.62 [95% CI 0.44–0.75]). Likewise, model 2 has a moderate correlation between CQ score and hippocampal volume (mean r in the fivefold cross validation, 0.61 [95% CI 0.43–0.74]). In the model 3, there was a high correlation between the adjusted CQ score and hippocampal volume (mean r in the fivefold cross-validation, 0.80 [95% CI 0.62–0.90]).

Assessment of the CQ score in the derivation cohort. In the derivation cohort ($n = 322$ subjects), the mean CQ score was -0.19 (standard deviation [SD], 2.72) in model 1 and 0.03 (SD, 0.67) in model 2. The mean hippocampal volume was 5480 mm^3 (SD, 1354 mm^3). In the derivation cohort, there was a moderate correlation between the CQ score and hippocampal volume in model 1 (mean r of the fivefold cross-validation, 0.62 [95% CI 0.44–0.75]; Fig. 1, model 1). Likewise, there was a moderate correlation between the CQ score and hippocampal volume in model 2 (mean r in the fivefold cross-validation, 0.61 [95% CI 0.43–0.74]; Fig. 1, model 2). In the model 3, there was a high correlation between the adjusted CQ score and hippocampal volume (mean r in the fivefold cross-validation, 0.80 [95% CI 0.62–0.90]; Fig. 1, model 3). In the graphical assessment, the CQ scores are widely distributed in the high score strata of the MMSE (Supplemental Fig. 2).

Validation of CQ score. In the validation cohort ($n = 96$ subjects), the mean CQ score was 2.81 (SD, 3.66) in model 1 and 0.54 (SD, 0.69) in model 2. The mean hippocampal volume was 6758 mm^3 (SD, 899 mm^3). In the validation cohort, there was a moderate correlation between the CQ score and hippocampal volume in model 1 ($r = 0.54$ [95% CI 0.38–0.67]; Fig. 2, model 1). Likewise, there was a moderate correlation between the CQ score and hippocampal volume in model 2 ($r = 0.53$ [95% CI 0.37–0.66]; Fig. 2, model 2). In the model 3, there was a high correlation between the adjusted CQ score and hippocampal volume (mean r in the fivefold cross-validation, 0.70 [95% CI 0.62–0.90]; Fig. 2, model 3). The detailed of the regression models are shown in Supplemental Table 2.

Discussion

In this study, there was a significant correlation between our newly developed CQ score based on five components of widely-used neuropsychological tests and the hippocampal volume. Furthermore, in the graphical assessment, the CQ scores were widely distributed in the high score strata of the MMSE, suggesting a ceiling effect of the MMSE consistent with our hypothesis that the CQ test could be a potential tool of screening individuals with small hippocampal volume. The strength of our study is the development of digital apps for cognitive testing that was studied with hippocampal volume using MRI.

Early detection of dementia facilitates early interventions that could prevent disease progression and maintain individuals at their current (time of examination) level of cognitive functions¹². Currently, there are various screening tools developed for early detection of dementia, including the MMSE, the Montreal Cognitive Assessment (MoCA) for mild cognitive impairment detection¹³, and the Clinical Dementia Rating scale (CDR)¹⁴. However, these tools have been known to possess their own limitations, including their long time taken, patient dependence, and requirement for human resources. The strengths of the CQ test are that it is brief (<5 min), self-explanatory and intuitive, easy to implement, and relies on objective measurements (i.e., digit forward, and digit backward, Stroop test, simple calculation, mental rotation test) based on the application on digital devices. While the correlation between the CQ score and hippocampal volume was moderate, a high accuracy might be a trade-off for simplicity. Additionally, the aim of the CQ test was not to diagnose dementia but to screen individuals who may have a small hippocampal volume and who may require further examinations (e.g., MRI) to detect early dementia.

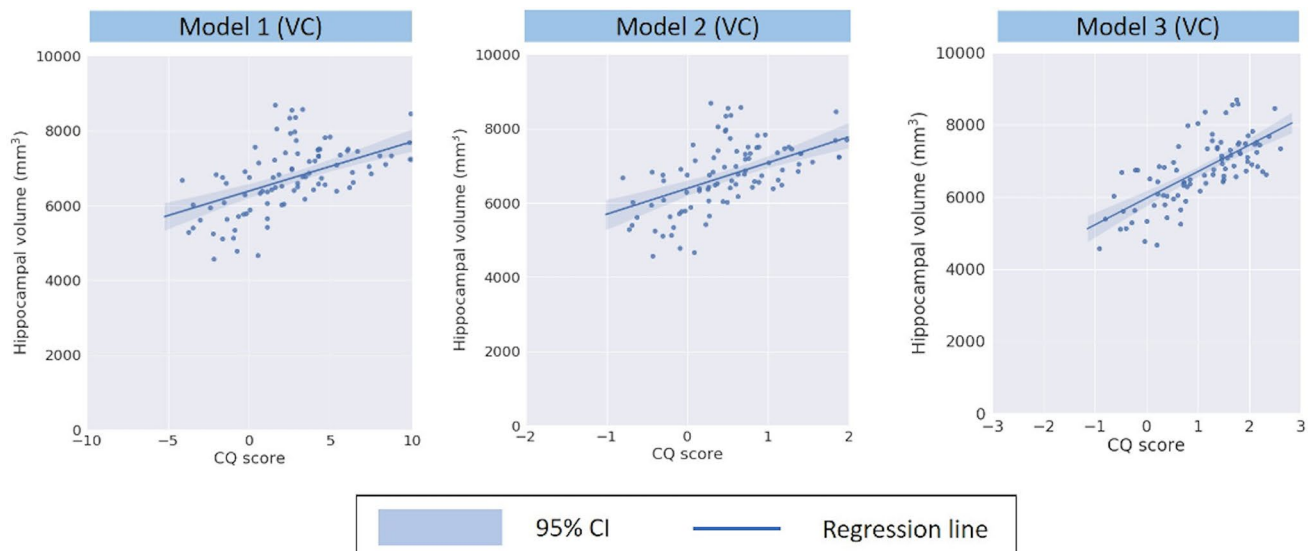


Figure 2. Correlation between Cognitive Quotient (CQ) score and hippocampal volume in the validation cohort. In the validation cohort, the model 1 has a moderate correlation between CQ score and hippocampal volume ($r=0.54$ [95% CI 0.38–0.67]). Likewise, the model 2 has moderate correlation between CQ score and hippocampal volume ($r=0.53$ [95% CI 0.37–0.66]). In the model 3, there was a high correlation between the adjusted CQ score and hippocampal volume (mean r in the fivefold cross-validation, 0.70 [95% CI 0.62–0.90]).

The conventional assessment tools represented by the MMSE and the MoCA have been criticized as having a ceiling effect in the very early stage of dementia or subclinical dementia^{15,16}. Instead, a small hippocampal volume has been reported as an early sign of dementia prior to clinical symptoms. Multiple studies have reported that a small hippocampal volume was found not only in 80–90% of patients with Alzheimer's disease^{17–20} but also in patients with mild cognitive impairment²¹, frontotemporal dementia²², and vascular dementia²³. In addition, a small hippocampal volume is associated with the risk of future Alzheimer disease²⁴. A 10-year follow-up study on 518 elderly patients found that a majority of the patients who developed dementia had a smaller baseline hippocampal volume years before their clinical diagnosis compared with those who remained dementia-free³. These findings collectively indicate that the subtle delayed memory decline with hippocampal volume atrophy can be observed long before a clinical diagnosis of dementia is made³. Despite the promising ability of hippocampal volume as a marker of early dementia, the hippocampal volume can only be evaluated using imaging examinations (e.g., MRI), which require financial costs and have limited availability. Thus, as a brief, inexpensive screening tool, the CQ test could be beneficial in identifying patients who need MRI to determine the presence of hippocampal atrophy that is not captured by the MMSE or other traditional screening tools that use the MMSE as a gold standard.

While formal validation is needed, by using the CQ test, physicians can easily screen people with seemingly normal cognitive function for early dementia without cost or time. In addition, because the CQ test is easy to use, people can repeat the test, and it is expected to be able to detect early cognitive decline from not only point estimates but also longitudinal changes of the test score. Lastly, in this study, we focused on the CQ test itself, while the prediction ability should be improved with the use of further characteristics information, such as smoking history, family history, comorbidities, and physical activities, with the use of machine-learning or artificial neural networks. The advantages of the application-based test is scalability; therefore, the development of CQ test as the first step should be an important basis for developing the optimal screening test for early dementia.

Potential limitations. Our study has several potential limitations. First, although there was a correlation between the CQ score and the hippocampal volume, it does not directly indicate early dementia. However, the primary objective of the CQ test is to screen individuals who need further investigations including high-cost examinations. Second, the study population included participants who underwent MRI for medical reasons, which might have affected cognitive function. However, the correlation between the CQ score and hippocampal volume remained significant in the cohort lacking any healthy participants (i.e., those that underwent MRI for medical check-up reasons). Third, patients in the validation cohort were assumed to be healthy population, there was no neuropsychological evaluation including MMSE. Nevertheless, there were no report on the cognitive problems and no specific findings in their MRI. Fourth, the proposed application-based test lacked the subtest that explicitly evaluates memory domain. This is because word recall, which is the widely used test to measure memory domain, is not well suited for an application designed to be completed in a short time. Thus, there may be room for improvement to the battery by adding an alternate subtest, such as FCSRT^{25,26}. Finally, there were no longitudinal data to evaluate the decline rate of the hippocampal volume because this study was based on a single measurement point.

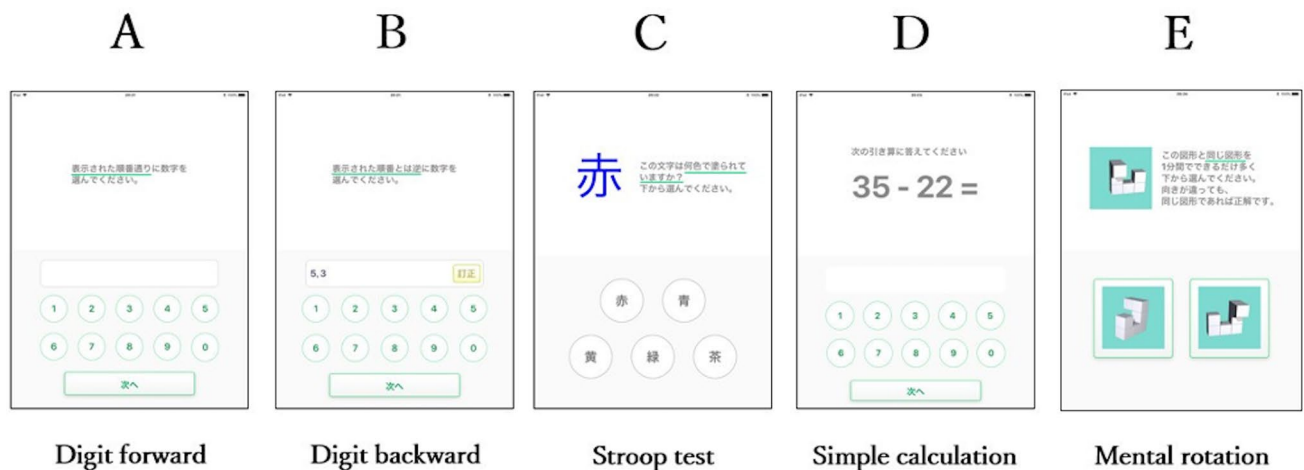


Figure 3. Application format and five questionnaires of the Cognitive Quotient (CQ) test. Panel (A) Digit forward test, Panel (B) Digit backward test, Panel (C) Stroop test, Panel (D) Simple calculation test, Panel (E) Mental rotation test.

Conclusions

In this analysis of 418 participants, our newly developed CQ score was significantly correlated with hippocampal volume. Our findings indicate that the CQ test could be a potential tool of screening individuals with small hippocampal volume, which is a marker of mild or subclinical cognitive impairment that cannot be detected by traditional screening tools. While further studies are warranted, this brief, inexpensive, application-based tool could be beneficial in identifying patients at-risk of cognitive problems that may warrant further investigation; not just for MRI but for other diagnostics, including a full neuropsychological evaluation, or other biomarkers.

Methods

Study design and setting. We performed an observational study of subjects who underwent brain MRI in seven institutions (four general hospitals and three clinics) from February 2018 to May 2019. There are approximately 30,000–320,000 annual outpatients in the seven institutions. This study was approved by the Institutional Review Board of Sado General Hospital. Written informed consents were obtained from all participants. This study on humans was carried out in accordance with Declaration of Helsinki.

Study participants. We included adult individuals (aged ≥ 18 years) who underwent brain MRI for a routine checkup or any medical reasons, such as a headache, dizziness, numbness, and medical follow up. When a physician decided to take a brain MRI, the physician also obtained written informed consent and performed a CQ test on the subject. The subject completed the CQ test before receiving the MRI findings. We excluded participants who were not independent, diagnosed with stroke, or whose hippocampal volume could not be successfully abstracted from the MRI results.

Development of the CQ test. We developed the CQ test as a screening tool for evaluating hippocampal volume to screen individuals who need further investigation for dementia. The CQ test consists of five components based on well-validated and widely-used neuropsychological tests^{27–34}. The five components are (1) digits forward (Fig. 3A), (2) digits backward (Fig. 3B), (3) Stroop test (Fig. 3C), (4) simple calculation (Fig. 3D), and (5) mental rotation (Fig. 3E).

Before conducting the current analysis, we conducted a pilot study to identify the optimal set of tests from September 2017 to November 2017. In the pilot study of 36 subjects, we used following test sets based on a priori knowledge: (1) digits forward, (2) digits backward, (3) Stroop test, (4) Addition, (5) Subtraction, (6) N-back task, and (7) delayed recall. Based on the results and correlations between tests, we integrated N-back, delayed recall tests into simple calculation test (i.e., addition and subtraction) because (1) N-back test without test volume limitation to avoid ceiling effect and delayed recall test were time-consuming and had limited feasibility for an application-based test, (2) N-back test was highly correlated with simple calculation ($r = 0.43$ with addition, $r = 0.16$ with subtraction), (3) and the correlation coefficient with the hippocampal volume decreased in the case of remaining test set including delayed recall (Supplemental Table 1). Thus, we first applied the following test: (1) digits forward, (2) digits backward, (3) Stroop test, (4) Simple calculation. By focusing on these four tests, testing time has been shortened, and therefore we have added mental rotation test according to the NIA-AA measurements. Consequently, the CQ test consists of (1) digits forward, (2) digits backward, (3) Stroop test, (4) Simple calculation, and (5) mental rotation, and these tests were consistent with the NIA-AA test except for vocabulary test (vocabulary test may be not feasible for an application-based test).

The first and second components are based on the widely-used cognitive test^{35,36}. The digits forward and digits backward tests are used to examine the function of working memory, which is related to the medial occipital cortex, dorsolateral prefrontal cortex, bilateral inferior parietal lobule, anterior cingulate, and Broca's area³⁷. For

example, there are significant positive correlations between the gray matter ratio, the percentage of gray matter volume in the intracranial volume, and performance on the Digit Span subtest³⁴. The third component, the Stroop test, is used to examine the cognitive regulation function and cognitive-related control of dorsolateral prefrontal and anterior cingulate cortex activity involving the hippocampus^{27,29,38}. This neuropsychological test is extensively used to assess the ability to inhibit cognitive interference that occurs when the processing of a specific stimulus feature impedes the simultaneous processing of a second stimulus attribute²⁷. The fourth component, simple calculation, is used to examine functions related to the bilateral premotor, posterior parietal, and prefrontal cortex³². A simple calculation is performed in assessment of prefrontal to posterior parietal cortex activation involving working memory^{33,39}. The fifth component, mental rotation of three-dimensional objects test, is used to examine hippocampal-related dorsal and ventral premotor cortex functions³⁰. Mental rotation has been described as a rotary transformation of a visual stimulus allowing it to be represented in a new orientation³¹. The CQ score can be calculated based on the results of these five questionnaires. We designed the CQ scoring system to evaluate where the total cognitive score is positioning on standard normal distribution of population. The total CQ score is derived as non-weighted (or weighted) sum of each test score also evaluated on normal distribution for each.

Main outcomes and measurements. The primary outcome measure was hippocampal volume (in mm³), which was measured by an automated-segmentation method from brain MRI as previously described (Supplemental Fig. 1)^{3,40}. We collected information of patient demographics, including age, sex, body mass index, smoking history, alcohol use, and the number of persons living with the subject. The MRI findings were reported by board-certified neurosurgeons, neurologists, or radiologists. The primary diagnosis at the MRI examination was abstracted from medical records when the subject had any disease or abnormality based on the physician's discretion.

Statistical analysis. We separated the data into derivation and validation cohorts. First, to develop the CQ scoring systems, we used data from six institutions (n = 322, 77% of the overall cohort) as the derivation cohort. The scoring systems were developed using following items: (1) for digit forward and backward: the maximum number of successfully-answered digits; (2) for simple calculation and the Stroop test: the number of correct answers and the mean time to answer, (3) for the mental rotation: the number of correct answers. Next, we built three scoring systems according to the previous literature⁴¹: (1) the sum of each test scores on standardized normal distribution (model 1), (2) multivariable linear regression model using each component (model 2), (3) and multivariable linear regression model including age and educational level. A is a set of 5 cognitive test set, and x is test score and μ and σ are mean and standard deviation for each. $z(x, \mu, \sigma)$ is a normal distribution of mean 0 and standard deviation 1 after the standardization.

Model 1

$$CQ = \sum_{i \in A} z(x_i, \mu_i, \sigma_i)$$

Model 2

$$CQ = \sum_{i \in A} a_i z(x_i, \mu_i, \sigma_i)$$

Model 3

$$CQ = \sum_{i \in A} a_i z(x_i, \mu_i, \sigma_i) + \sum_{j \in \{age, education\}} b_j z(x_j, \mu_j, \sigma_j)$$

After developing these scoring systems, we used fivefold cross-validation of the derivation cohort to examine the models' performance of the models. In the validation cohort using data from a clinic (n = 96, 23% of the overall cohort), we examined the association between the developed CQ scoring systems and the actual hippocampal volume. In the validation cohort, all subjects had undergone MRI for a medical checkup without any symptoms (i.e., healthy population), since the development of our screening tool aimed to identify early dementia that was not identified in the conventional screening tools due to ceiling effect. In addition, we also graphically-assessed the distribution and association between the CQ score and MMSE by using scatter-plot in the derivation cohort. For MRI data, we used Hippodeep⁴⁰ on Mindboggle⁴² as an alternative tool of FreeSurfer⁴³ which takes over 10 h for pre-processing to extract the hippocampal volume. Hippodeep is a python-based hippocampal region extraction tool and its processing time is very short (less than 1 min). Moreover, we analyzed data using python 3.5.6 and SciPy library (1.0.0) on Google Datalab (vCPU × 1, 3.75 GB Memory) platform.

Data availability

The datasets generated during and/or analyzed during the current study are not publicly available due to no IRB's approval of data sharing.

Received: 17 March 2020; Accepted: 30 July 2020

Published online: 07 October 2020

References

1. Ferri, C. P. *et al.* Global prevalence of dementia: A Delphi consensus study. *Lancet* **366**, 2112–2117 (2005).
2. Burns, A. & Iliffe, S. Dementia. *BMJ* **338**, b75 (2009).

3. den Heijer, T. *et al.* A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain* **133**, 1163–1172 (2010).
4. den Heijer, T. *et al.* Use of hippocampal and amygdalar volumes on magnetic resonance imaging to predict dementia in cognitively intact elderly people. *Arch Gen Psychiatry* **63**, 57–62 (2006).
5. Matsuda, H. *et al.* Automatic voxel-based morphometry of structural MRI by SPM8 plus diffeomorphic anatomic registration through exponentiated lie algebra improves the diagnosis of probable Alzheimer Disease. *AJNR Am. J. Neuroradiol.* **33**, 1109–1114 (2012).
6. Robinson, L., Tang, E. & Taylor, J. P. Dementia: timely diagnosis and early intervention. *BMJ* **350**, h3029 (2015).
7. Pendlebury, S. T., Cuthbertson, F. C., Welch, S. J., Mehta, Z. & Rothwell, P. M. Underestimation of cognitive impairment by mini-mental state examination versus the montreal cognitive assessment in patients with transient ischemic attack and stroke: A population-based study. *Stroke* **41**, 1290–1293 (2010).
8. Jack, C. R. Jr. *et al.* NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* **14**, 535–562 (2018).
9. Vijayakumar, A. & Vijayakumar, A. Comparison of hippocampal volume in dementia subtypes. *ISRN Radiol.* **2**, 174524 (2013).
10. Dolek, N., Saylisoy, S., Ozbabalik, D. & Adapinar, B. Comparison of hippocampal volume measured using magnetic resonance imaging in Alzheimer's disease, vascular dementia, mild cognitive impairment and pseudodementia. *J. Int. Med. Res.* **40**, 717–725 (2012).
11. Trzepacz, P. T. *et al.* Relationship between the Montreal cognitive assessment and mini-mental state examination for assessment of mild cognitive impairment in older adults. *BMC Geriatr.* **15**, 107 (2015).
12. Boustani, M., Peterson, B., Hanson, L., Harris, R. & Lohr, K. N. Force USPST Screening for dementia in primary care: A summary of the evidence for the U.S Preventive Services Task Force. *Ann. Intern. Med.* **138**, 927–937 (2003).
13. Nasreddine, Z. S. *et al.* The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).
14. Morris, J. C. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
15. Mitchell, A. J. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J. Psychiatr. Res.* **43**, 411–431 (2009).
16. Lawrence, J., Davidoff, D., Katt-Lloyd, D., Auerbach, M. & Hennen, J. A pilot program of improved methods for community-based screening for dementia. *Am. J. Geriatr. Psychiatry* **9**, 205–211 (2001).
17. Ridha, B. H. *et al.* Application of automated medial temporal lobe atrophy scale to Alzheimer disease. *Arch. Neurol.* **64**, 849–854 (2007).
18. Colliot, O. *et al.* Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology* **248**, 194–201 (2008).
19. Fox, N. C. *et al.* Presymptomatic hippocampal atrophy in Alzheimer's disease. A longitudinal MRI study. *Brain* **119**, 2001–2007 (1996).
20. Fox, N. C., Warrington, E. K., Stevens, J. M. & Rossor, M. N. Atrophy of the hippocampal formation in early familial Alzheimer's disease. A longitudinal MRI study of at-risk members of a family with an amyloid precursor protein 717Val-Gly mutation. *Ann. N. Y. Acad. Sci.* **777**, 226–232 (1996).
21. Devanand, D. P. *et al.* Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. *Neurology* **68**, 828–836 (2007).
22. Barnes, J. *et al.* Measurements of the amygdala and hippocampus in pathologically confirmed Alzheimer disease and frontotemporal lobar degeneration. *Arch. Neurol.* **63**, 1434–1439 (2006).
23. Kril, J. J., Patel, S., Harding, A. J. & Halliday, G. M. Patients with vascular dementia due to microvascular pathology have significant hippocampal neuronal loss. *J. Neurol. Neurosurg. Psychiatry* **72**, 747–751 (2002).
24. Du, A. T. *et al.* Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiol. Aging* **27**, 733–740 (2006).
25. Grober, E. *et al.* Memory impairment on free and cued selective reminding predicts dementia. *Neurology* **54**, 827–832 (2000).
26. Donohue, M. C. *et al.* The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol.* **71**, 961–970 (2014).
27. Scarpina, F. & Tagini, S. The stroop color and word test. *Front. Psychol.* **8**, 557 (2017).
28. Cohen, J. D. *et al.* Temporal dynamics of brain activation during a working memory task. *Nature* **386**, 604–608 (1997).
29. MacDonald, A. W. 3rd., Cohen, J. D., Stenger, V. A. & Carter, C. S. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* **288**, 1835–1838 (2000).
30. Shepard, R. N. & Metzler, J. Mental rotation of three-dimensional objects. *Science* **171**, 701–703 (1971).
31. Cona, G., Panozzo, G. & Semenza, C. The role of dorsal premotor cortex in mental rotation: A transcranial magnetic stimulation study. *Brain Cogn.* **116**, 71–78 (2017).
32. Rueckert, L. *et al.* Visualizing cortical activation during mental calculation with functional MRI. *Neuroimage* **3**, 97–103 (1996).
33. Cowell, S. F., Egan, G. F., Code, C., Harasty, J. & Watson, J. D. The functional neuroanatomy of simple calculation and number repetition: A parametric PET activation study. *Neuroimage* **12**, 565–573 (2000).
34. Gerald G. WAIS-III in *Wechsler Adult Intelligence Scale (3rd ed.)*/WMS-III: *Wechsler Memory Scale (3rd ed.) Technical Manual.* (Psychological Corporation, 1997).
35. Hartman, D. E. Wechsler Adult Intelligence Scale IV (WAIS IV): return of the gold standard. *Appl. Neuropsychol.* **16**, 85–87 (2009).
36. Taki, Y. *et al.* Correlation between gray/white matter volume and cognition in healthy elderly people. *Brain Cogn.* **75**, 170–176 (2011).
37. Gerton, B. K. *et al.* Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. *Neuropsychologia* **42**, 1781–1787 (2004).
38. Khan, S. A. *et al.* The hippocampus and executive functions in depression. *Ind. Psychiatry J.* **24**, 18–22 (2015).
39. Baddeley, A., Jarrold, C. & Vargha-Khadem, F. Working memory and the hippocampus. *J. Cogn. Neurosci.* **23**, 3855–3861 (2011).
40. Thyreau, B., Sato, K., Fukuda, H. & Taki, Y. Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Med. Image Anal.* **43**, 214–228 (2018).
41. Chandler, M. J. *et al.* A total score for the CERAD neuropsychological battery. *Neurology* **65**, 102–106 (2005).
42. Klein, A. *et al.* Mindboggling morphometry of human brains. *PLoS Comput Biol.* **13**, e1005350 (2017).
43. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).

Acknowledgements

The authors would like to thank Dr. Kazuhiro Sampei, Department of Neurology, Sado General Hospital for supporting study participants. This work was supported by a Grant from the Ministry of Internal Affairs and Communications of Japan: I-Challenge!, ICT innovation challenge program.

Author contributions

W.K. takes responsibility for the paper as a whole. W.K., Y.A., and K.S., conceived the study. T.G., Y.A., and K.S. supervised the conduct of the study. T.G. provided statistical advice. W.K. analyzed the data. W.K. drafted the manuscript, and all authors contributed substantially to its revision.

Competing interests:

Mr. Kasai is an employee of Splink. Dr. Goto has performed clinical-research related consultation for Splink. Mr. Aoyama is the Chief Executive Officer (CEO) of Splink. Dr. Sato does not have any conflict of interest.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-74019-7>.

Correspondence and requests for materials should be addressed to T.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020