Early identification and heritability of mild cognitive impairment

William S Kremen,1,2* Amy J Jak,1,2,3 Matthew S Panizzon,1 Kelly M Spoon,1,4 Carol E Franz,1 Wesley K Thompson,1 Kristen C Jacobson,5 Terrie Vasilopoulou,5 Eero Vuoksimaa,1,6 Hong Xian,7,8 Rosemary Toomey9 and Michael J Lyons9

1Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA, 2Center of Excellence for Stress and Mental Health, VA San Diego Healthcare System, La Jolla, CA, USA, 3Psychology Service, VA San Diego Healthcare System, La Jolla, CA, USA, 4Computational Science Research Center, San Diego State University/Claremont Graduate University, San Diego, CA, USA, 5Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA, 6Department of Public Health, University of Helsinki, Helsinki, Finland, 7Department of Statistics, St Louis University, St Louis, MO, USA, 8Research Service, VA St Louis Healthcare System, St Louis, MO, USA and 9Department of Psychology, Boston University, Boston, MA, USA

*Corresponding author. Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive (MC 0738), La Jolla, CA 92037, USA. E-mail: wkremen@ucsd.edu

Accepted 4 November 2013

Background  Identifying mild cognitive impairment (MCI) in midlife could improve early identification of Alzheimer’s disease (AD). Also, AD is highly heritable, but the heritability of MCI has not been established. We estimated prevalence rates, association with premorbid general cognitive ability (GCA) and heritability for different definitions of neuropsychologically defined MCI in adults in their 50s.

Method  We examined 1126 twins aged 51–59 years when recruited into the Vietnam Era Twin Study of Aging (VETSA). Six neurocognitive domains were assessed using tests designed to avoid ceiling effects. To differentiate MCI from low overall ability, criteria included adjustment for GCA measured at approximately age 20 years.

Results  As in older adults, prevalence rates varied widely. Among the lower prevalence rates were some definitions of multiple-domain MCI and single-domain amnestic MCI, which may be less likely than other MCI categories to revert to normal on follow-up. Low prevalence rates in middle-aged adults are also more likely to be valid. MCI was also associated with lower premorbid GCA. Heritability estimates for any MCI and amnestic MCI averaged .40–.48.

Conclusions  By testing multiple cognitive domains and avoiding ceiling effects, MCI can be identified before age 60 years. Premorbid GCA is a risk/protection factor, but deficits after adjusting for early adult GCA suggest additional processes leading to declining trajectories. Heritabilities were comparable to AD, suggesting MCI as an appropriate phenotype for genetic association studies. Full validation will require follow-up assessments (currently under way). Community-based studies are important for this early identification because adults of this age are unlikely to present in clinics.

Keywords  Cognitive decline, Alzheimer’s disease, preclinical diagnosis, risk factor, middle age

Published by Oxford University Press on behalf of the International Epidemiological Association
Introduction

There is growing emphasis on early identification of Alzheimer’s disease (AD).1–4 The pathophysiological process begins 10–15 years before onset of dementia.5–9 Yet studies of middle-aged adults are lacking. Mild cognitive impairment (MCI) does not always develop into AD10,11 but it is central to early identification. Average age in most MCI studies is in the 70s.12,13 Prevalence rates of 0%–13.7% have been reported in adults under 65 years of age, but sample sizes were small.14–17 Shifting the focus to MCI in midlife would enhance the potential to shed light on early trajectories leading toward different types of cognitive decline or maintenance of function. Here we investigated early identification of MCI using a genetic epidemiological approach.

Testing has often inadequately covered cognitive abilities (e.g., omitting visual-spatial memory18) and tests designed for older samples are susceptible to ceiling effects in younger adults.19 Defining abnormality based on impairment on a single test results in excess false positives,20–25 yet that remains common for MCI. There is also no agreement on cut-points for impairment.25,26

In the largest extant study, heritability—proportion of phenotypic variance due to genes—of AD was .58.26 If at least some MCI is related to AD, then MCI ought to be heritable. We are aware of two studies, both with negative findings for heritability of MCI-related conditions. The authors of one acknowledged limitations due to small sample size.27 The other had a very large sample, necessitating a brief telephone cognitive assessment that could have lacked sufficient sensitivity.28 Thus, the heritability of MCI remains uncertain.

We delineated five criteria sets for neuropsychologically-defined MCI using tests designed to cover a range of cognitive functions and cut-offs for cognitive impairment, and to avoid ceiling effects in a community-dwelling, middle-aged sample. We capitalized on having general cognitive ability (GCA) scores when participants were approximately 20 years old, roughly 35 years prior to our midlife assessment. We compared prevalence rates of MCI for each of the different operational definitions based on scores adjusted for age 20 GCA. We then tested whether premorbid GCA was lower in MCI participants (before adjustment). We also estimated the magnitude of genetic and environmental influences on MCI according to each definition (based on adjustment for age 20 GCA) to determine heritabilities. Finally, because APOE-ε4 is a well-established risk factor for AD,29 we tested whether the proportion of ε4 carriers was higher in MCI than in non-MCI groups.

Methods

Participants
Participants were in wave 1 of the Vietnam Era Twin Study of Aging (VETSA).30,31 Participants were in military service sometime between 1965 and 1975. Most (~80%) reported no combat experience. Some 35 years later, they constituted a relatively representative sample of middle-aged men living throughout the USA, based on census data for health and demographic characteristics32 (Table 1). There were 1237 participants (349 monozygotic (MZ) pairs, 265 dizygotic (DZ) pairs, 9 unpaired twins). Participants underwent assessments at the University of California, San Diego or Boston University.

Zygosity was determined by 25 microsatellite markers; for 8% of the sample, zygosity was determined by questionnaires and blood group, with 95% agreement with DNA-based results. Exclusions for MCI analyses were: medical conditions that might cause non-MCI-related cognitive deficits: seizure disorder; multiple sclerosis; stroke; HIV/AIDS; schizophrenia; serious alcohol dependence; brain cancer; or dementia. Depressive symptoms and traumatic brain injury (TBI) were not exclusion criteria because they are risk factors for dementia. We added questions about head injury after the study began, so only 1023 participants were asked the question used in the Epidemiological Catchment Area (ECA) study: ‘Have you ever had a severe head injury that was associated with a loss of consciousness or confusion?’33 There were 110 (10.75%) VETSA participants who answered ‘yes’, compared with 10.86% of men in the ECA study. Depressive symptoms were assessed with the Center for Epidemiological Studies Depression Scale (CES-D).34 The mean score was 8.31 (standard deviation (SD) = 8.25), and there were 185/1237 (14.96%) individuals at the cut-off score of ≥16 for major depressive disorder.30 In three mixed-sex community samples in the original CES-D report, the weighted mean score was 8.65 and the weighted mean percentage of individuals with scores ≥16 was 17.01%. Moreover, depressive symptoms and cognition shared ≤1% of variance in the VETSA sample, an effect comparable to other studies of nonpatient samples.30 Thus, the present sample did not have elevated rates of TBI or depression.

Although alcohol use disorders are also a risk factor for dementia, we did exclude what we considered to be very serious alcohol dependence because we were concerned that the effects of very severe alcohol dependence might substantially overshadow other effects. We excluded individuals who responded ‘yes’ to the question: ‘Have you ever been told by a doctor that you had alcohol abuse or dependency?’29 We reasoned that only very severe cases would have been told this by their doctor, and the low rate (43 individuals, 3.48%) is consistent with that notion.

Neurocognitive assessment
GCA was assessed with the Armed Forces Qualification Test (AFQT); AFQT is highly correlated (r = .84) with standard IQ measures.35 It was administered to military induction at an average age of
Classifications were further organized into subtypes: amnestic (A-MCI); non-amnestic (NA-MCI); single-domain; multiple-domain; any (any-MCI).

**Composite scores.** In Ganguli et al.’s approach, domain composite scores are the mean of Z-scored within a domain. We examined impairment cut-points at the 5th and 2.5th percentiles, corresponding closely to 1.65 and 2 SDs below the mean, respectively. These are referred to as ‘Composite 5’ and ‘Composite 2.5.’

**Number-of-measures impaired.** Jak et al. defined a domain as impaired based on number of measures below a cut-point. ‘Typical criteria’ define impairment by performance on one measure greater than 1.5 SDs below the mean; this is similar to Petersen’s criteria, probably the most commonly used metric. ‘Comprehensive criteria’ relax the cut-point to 1 SD, but require performance on at least two measures to be below that cut-point. The ‘Conservative criteria’ require at least two impaired measures, but with a more conservative cut-point of 1.5 SDs below the mean.

**Statistical analysis**

Missing neuropsychological values (<1%) were replaced using predictive mean matching with multivariate imputation by chained equations (MICE package version 2.10 in R version 2.13.1) Missing age 20 AFQT scores (n = 15) were not imputed. After all exclusions, there were 1126 remaining participants.

We controlled for early adult AFQT scores to provide an objective index of change, thereby ensuring that MCI classifications represented declines from previous levels of functioning, and not just the lower end of the ability distribution. For example, scoring below the 5th percentile means that those individuals were below that level after adjusting for age 20 GCA.

To further understand the role of premorbid GCA, we utilized MCI classifications that were not adjusted for age 20 AFQT. We compared age 20 AFQT scores in non-MCI, single-domain and multiple-domain MCI. These comparisons were based on mixed linear models that adjust for non-independence of observations (twins within pairs). Neuropsychological scores adjusted for age 20 AFQT could not be looked up in standard tables of age- and education-adjusted values. Consequently, for all analyses we used the VETSA sample as our normative sample so we could compare prevalence rates based on MCI adjusted and unadjusted for early GCA. As noted, the VETSA sample is reasonably representative of men in their age range, and the overall sample of 1237 is larger than many normative samples for neuropsychological tests.

Twin analyses to estimate MCI heritability were performed via maximum-likelihood-based methods using Open Mx. The classical twin design divides phenotypic variance into: additive genetic (A); common

---

**Table 1 Sample characteristics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (SD)</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>55.4 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Education (years)</td>
<td>13.8 (2.1)</td>
<td></td>
</tr>
<tr>
<td>AFQT score at average age 20</td>
<td>61.1 (22.3)</td>
<td></td>
</tr>
<tr>
<td>AFQT score at average age 55</td>
<td>64.1 (20.9)</td>
<td></td>
</tr>
<tr>
<td>CES-D depressive symptoms</td>
<td>8.31 (8.25)</td>
<td>N (3.9 %)</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unskilled worker</td>
<td>21 (2.8%)</td>
<td></td>
</tr>
<tr>
<td>Semi-skilled, skilled manual</td>
<td>455 (36.8%)</td>
<td></td>
</tr>
<tr>
<td>Technical, clerical, small business, minor professional</td>
<td>677 (54.8%)</td>
<td></td>
</tr>
<tr>
<td>Major professional, large business</td>
<td>69 (5.6%)</td>
<td></td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>952 (77.7%)</td>
<td></td>
</tr>
<tr>
<td>Part-time</td>
<td>249 (5.8%)</td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>111 (9.1%)</td>
<td></td>
</tr>
<tr>
<td>Other (disabled, unemployed, student)</td>
<td>91 (7.4%)</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>968 (78.7%)</td>
<td></td>
</tr>
<tr>
<td>Separated/divorced</td>
<td>178 (14.5%)</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>18 (1.5%)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>66 (5.3%)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>1052 (86%)</td>
<td></td>
</tr>
<tr>
<td>Health status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of hypertension</td>
<td>689 (55.7%)</td>
<td></td>
</tr>
<tr>
<td>History of diabetes</td>
<td>104 (8.4%)</td>
<td></td>
</tr>
<tr>
<td>History of cardiovascular disease</td>
<td>214 (17.3%)</td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>296 (24%)</td>
<td></td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never or not in past 2 weeks</td>
<td>500 (40.4%)</td>
<td></td>
</tr>
<tr>
<td>1 or fewer drinks per day</td>
<td>504 (40.7%)</td>
<td></td>
</tr>
<tr>
<td>&gt;1 and ≤2 drinks per day</td>
<td>105 (10.4%)</td>
<td></td>
</tr>
<tr>
<td>&gt;2 drinks per day</td>
<td>134 (10.8%)</td>
<td></td>
</tr>
</tbody>
</table>

AFQT, Armed Forces Qualification Test; CES-D, Center for Epidemiological Studies Depression Scale (16 is a standard threshold for clinical depression).
environmental (C); nonadditive/dominance genetic (D); and unique environmental (E) influences. C refers to environmental factors that make twins similar; E refers to environmental factors that make twins different, including measurement error. MZ twins generally share 100% of their genes and DZ twins share an average of 50% of their segregating genes. Additive genetic factors are assumed to correlate 1.0 in MZ twins and .50 in DZ twins. Nonadditive (interactive) genetic factors are assumed to correlate .25 in DZ twins because there is a 25% chance that DZ twins share both alleles at a given locus. C is assumed to correlate 1.0 for both types of twins. E is, by definition, uncorrelated between members of a twin pair. An ACDE model cannot be identified, so ACE and ADE models are tested separately. MZ correlations substantially more than twice the DZ correlations indicate that ADE models are most appropriate.

We fit univariate ACE or ADE models to the data with MCI as a 3-level ordinal variable [non-MCI (cognitively normal), single-domain MCI, multiple-domain MCI] and a binary variable (non-MCI vs MCI). Model fits were tested relative to a saturated model based on the likelihood-ratio chi-square test (LRT), which is the difference in the $-2 \log$ likelihood ($-2 \text{LL}$) of the hypothesized model from that of the saturated model. Nonsignificant LRTs indicate good model fit (i.e. nonsignificant reduction in fit relative to the saturated model). When C accounted for $\leq 10\%$ of the variance, we also tested reduced AE models. For ADE models, we always tested AE models because a substantially larger sample is needed to differentiate A and D effects. Model comparisons were based on Akaike’s Information Criterion (AIC). Smaller AICs indicate a preferred balance between model fit and number of parameters.

Results

MCI prevalence rates

For single-domain and multiple-domain MCI, rates were 1.24–29.93% and 1.33–34.81%, respectively. Prevalence rates were relatively low (≤10%) for all categories of multiple-domain MCI for the Composite 2.5, Composite 5 and Conservative criteria. For single-domain MCI, they were low for all A-MCI criteria except the Comprehensive criteria, and for the NA-MCI and any-MCI Composite 2.5 criteria (Table 4). Definitions with low prevalence rates of

<table>
<thead>
<tr>
<th>Cognitive domain</th>
<th>Tests and measures</th>
<th>No. of measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episodic Memory</td>
<td>CVLT-2: Sum of trials 1-5; delayed free recall</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>WMS-3: Logical memories immediate, delayed free recall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WMS-3: Visual reproduction immediate and delayed free recall</td>
<td></td>
</tr>
<tr>
<td>Executive Function</td>
<td>DKEFS Trails: Switching (condition 4)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>DKEFS Fluency: Category switching</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroop: Color-word, interference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WASI: Matrix reasoning</td>
<td></td>
</tr>
<tr>
<td>Attention/Working Memory</td>
<td>WMS-3: Digit span</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>WMS-3: Spatial span</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WMS-3: Letter-number sequencing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DKEFS Trails: Cancellations (condition 1)</td>
<td></td>
</tr>
<tr>
<td>Verbal/Language</td>
<td>WASI: Vocabulary</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>DKEFS: Letter fluency</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DKEFS: Category fluency</td>
<td></td>
</tr>
<tr>
<td>Visual-Spatial</td>
<td>Gottschaldt Hidden Figures</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Card rotation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WMS-3: Visual reproduction copy</td>
<td></td>
</tr>
<tr>
<td>Processing Speed</td>
<td>DKEFS Trails: Numbers (condition 2)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>DKEFS Trails: Letters (condition 3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroop: Word</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroop: Color</td>
<td></td>
</tr>
</tbody>
</table>

MCI, mild cognitive impairment; CVLT-2, California Verbal Learning Test-Version 2; WMS-3, Wechsler Memory Scale-Version 3; DKEFS, Delis-Kaplan Executive Function System; WASI, Wechsler Abbreviated Scale of Intelligence.

*Standard WMS-3 instructions call for reading the second Logical Memory story a second time, but it was read only once in our administration.
course, have large majorities classified as cognitively normal, which one would expect in this age range. Prevalence rates broken down according to the particular domain impaired are shown in Supplementary Table 2 (available as Supplementary data at IJE online); there was a slight tendency for higher prevalence for the episodic memory (amnestic) and executive domains.

As expected, fewer participants were classified as MCI when premorbid GCA was taken into account. For example, rates for adjusted classifications of the three-level any-MCI measure (shown first) vs unadjusted classifications were as follows: Typical (67.74% vs 73.28%); Comprehensive (56.66% vs 61.88%); Conservative (24.96% vs 34.88%).

Relationship of MCI to early adult GCA

For classifications not adjusted for premorbid GCA, age 20 AFQT scores were lowest in multiple-domain MCI, followed by single-domain, and then non-MCI participants (Supplementary Table 3, available as Supplementary data at IJE online).

### Table 3 Summary of MCI definitions

<table>
<thead>
<tr>
<th>MCI definition</th>
<th>Impairment cut-point based on normative values</th>
<th>No. of impaired measures required within a domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite score method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>5th percentile</td>
<td>Average of all b</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>2.5th percentile</td>
<td>Average of all b</td>
</tr>
<tr>
<td>No. measures impaired method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>1.5 SDs</td>
<td>1</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>1 SD</td>
<td>2</td>
</tr>
<tr>
<td>Conservative</td>
<td>1.5 SDs</td>
<td>2</td>
</tr>
</tbody>
</table>

MCI, mild cognitive impairment.

*Each definition can be subclassified according to single-domain vs multi-domain, and A-MCI vs any-MCI. All definitions were based on classification after adjusting for early adult cognitive ability.

bAverages were computed after \( z \)-scoring each measure.

### Table 4 MCI prevalence rates and polychoric and tetrachoric correlations for monozygotic (MZ) and dizygotic (DZ) twins

<table>
<thead>
<tr>
<th>MCI definition</th>
<th>MCI prevalence</th>
<th>Twin correlations three-level measure a</th>
<th>Twin correlations binary measure b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single domain</td>
<td>Multiple domain</td>
<td>( r_{MZ} )</td>
</tr>
<tr>
<td>Any MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>15.54%</td>
<td>5.86%</td>
<td>.64 (.49; .76)</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>7.73%</td>
<td>3.20%</td>
<td>.50 (.24; .70)</td>
</tr>
<tr>
<td>Typical</td>
<td>29.93%</td>
<td>34.81%</td>
<td>.43 (.30; .55)</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>29.40%</td>
<td>27.26%</td>
<td>.52 (.40; .63)</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>18.92%</td>
<td>6.04%</td>
<td>.42 (.24; .58)</td>
</tr>
<tr>
<td>Amnestic MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>3.02%</td>
<td>2.04%</td>
<td>–</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>1.24%</td>
<td>1.33%</td>
<td>–</td>
</tr>
<tr>
<td>Typical</td>
<td>7.55%</td>
<td>16.87%</td>
<td>.43 (.24; .59)</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>10.75%</td>
<td>15.99%</td>
<td>.56 (.40; .69)</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>7.64%</td>
<td>3.29%</td>
<td>–</td>
</tr>
<tr>
<td>Non-amnestic MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>12.52%</td>
<td>3.82%</td>
<td>–</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>6.49%</td>
<td>1.87%</td>
<td>–</td>
</tr>
<tr>
<td>Typical</td>
<td>22.38%</td>
<td>17.94%</td>
<td>.27 (.11; .41)</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>18.65%</td>
<td>11.27%</td>
<td>.29 (.11; .45)</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>11.28%</td>
<td>2.75%</td>
<td>–</td>
</tr>
</tbody>
</table>

MCI, mild cognitive impairment; \( S+M \), combination of single- or multiple-domain MCI. All MCI definitions were based on classification after adjusting for early adult cognitive ability. Correlations shown in bold had 95% confidence intervals that did not include 0. Dashed lines are shown where contingency tables for MCI definitions contained some cells with too few participants to compute meaningful correlations.

aThree-level measure: no MCI, single-domain MCI, multiple-domain MCI.

bBinary measure: no MCI, single-domain or multiple-domain MCI.
MCI heritability

Table 4 shows the MZ and DZ correlations for different MCI definitions. For the three-level measure, meaningful correlations could not be computed for some definitions of A-MCI and NA-MCI because some cells in the contingency tables of twin pair concordance contained too few participants. Collapsing single- and multiple-domain categories for the binary definition allowed for calculation of correlations for all but one MCI definition. MZ correlations were higher than DZ correlations for most definitions, suggesting the importance of genetic influences. The three number-of-measures-impaired definitions for NA-MCI had higher DZ than MZ correlations, most likely reflecting unreliability of estimates due to their low prevalence rates.

Model-fitting results for the three-level measure are shown in Table 5. In the best-fitting model for any-MCI, heritabilities (‘A’ in Table 5) were .37–.63. For all but the Comprehensive criteria, the full models were ADE models. C influences tended to be very low, and could be dropped without significant reductions in fit. Given the correlations in Table 4, models for A-MCI and NA-MCI could only be tested for the Typical and Comprehensive criteria. In those models, C accounted for 12–39% of the variance, and heritability appeared to be higher for A-MCI than for NA-MCI; however, each of the heritabilities had confidence intervals that included zero.

Model-fitting results for binary MCI definitions are shown in Table 6. Except for the Conservative criteria, heritability estimates for any-MCI were similar to those for the three-level definitions (.21–.64). MZ and DZ correlations for binary and three-level Conservative criteria did not vary dramatically (Table 4). However, heritabilities differed because each correlation changed a relatively small amount but in opposite directions. For binary A-MCI definitions, heritabilities were .22–.56; however, only estimates for the Comprehensive and Conservative criteria were significant. Other estimates suggested moderate heritability, but confidence intervals were wide and included zero. For binary NA-MCI, both composite score measures were significantly heritable (.59 and .47). The best-fitting models for all of the number-of-measures-impaired definitions were CE models with heritability estimates of zero.

Table 5 Model-fitting results for three-level MCI definitions adjusted for cognitive ability in early adulthood

<table>
<thead>
<tr>
<th>MCI definition</th>
<th>Model</th>
<th>A (95% CI)</th>
<th>C (95% CI)</th>
<th>D (95% CI)</th>
<th>E (95% CI)</th>
<th>-2LL</th>
<th>P</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any MCI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>ADE</td>
<td>.39 (.00; .74)</td>
<td>–</td>
<td>.26 (.00; .75)</td>
<td>.35 (.24; .50)</td>
<td>1398.7</td>
<td>0.305</td>
<td>–845.31</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.63 (.49; .75)</td>
<td>–</td>
<td>–</td>
<td>.37 (.25; .51)</td>
<td>1399.0</td>
<td>0.378</td>
<td>–846.99</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>ADE</td>
<td>.23 (.00; .69)</td>
<td>–</td>
<td>.29 (.00; .71)</td>
<td>.48 (.29; .74)</td>
<td>910.1</td>
<td>0.920</td>
<td>–1333.86</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.50 (.26; .69)</td>
<td>–</td>
<td>–</td>
<td>.50 (.31; .74)</td>
<td>910.4</td>
<td>0.948</td>
<td>–1335.65</td>
</tr>
<tr>
<td>Typical</td>
<td>ADE</td>
<td>.12 (.00; .51)</td>
<td>–</td>
<td>.30 (.00; .54)</td>
<td>.58 (.46; .71)</td>
<td>2432.1</td>
<td>0.066</td>
<td>188.14</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.40 (.28; .52)</td>
<td>–</td>
<td>–</td>
<td>.60 (.48; .72)</td>
<td>2432.9</td>
<td>0.084</td>
<td>186.85</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.43 (.04; .61)</td>
<td>.08 (.00; .40)</td>
<td>–</td>
<td>.49 (.39; .61)</td>
<td>2361.5</td>
<td>0.309</td>
<td>117.46</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.52 (.40; .62)</td>
<td>–</td>
<td>–</td>
<td>.48 (.38; .60)</td>
<td>2361.7</td>
<td>0.396</td>
<td>115.65</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>ADE</td>
<td>.00 (.00; .49)</td>
<td>–</td>
<td>.40 (.00; .56)</td>
<td>.60 (.45; .77)</td>
<td>1556.9</td>
<td>0.328</td>
<td>–687.14</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.37 (.20; .52)</td>
<td>–</td>
<td>–</td>
<td>.63 (.48; .80)</td>
<td>1558.4</td>
<td>0.298</td>
<td>–687.65</td>
</tr>
<tr>
<td>Amnestic MCI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>ACE</td>
<td>.19 (.00; .59)</td>
<td>.24 (.00; .50)</td>
<td>–</td>
<td>.57 (.41; .75)</td>
<td>1565.5</td>
<td>0.234</td>
<td>–678.47</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.43 (.00; .67)</td>
<td>.12 (.00; .52)</td>
<td>–</td>
<td>.46 (.33; .61)</td>
<td>1667.1</td>
<td>0.165</td>
<td>–576.89</td>
</tr>
<tr>
<td>Non-amnestic MCI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>ACE</td>
<td>.00 (.00; .00)</td>
<td>.28 (.00; .39)</td>
<td>–</td>
<td>.72 (.58; .84)</td>
<td>2121.4</td>
<td>0.821</td>
<td>–122.64</td>
</tr>
<tr>
<td></td>
<td>CE</td>
<td>–</td>
<td>.28 (.16; .39)</td>
<td>–</td>
<td>.72 (.61; .84)</td>
<td>2121.4</td>
<td>.894</td>
<td>–124.64</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.12 (.00; .46)</td>
<td>.18 (.00; .39)</td>
<td>–</td>
<td>.71 (.54; .86)</td>
<td>1805.4</td>
<td>0.852</td>
<td>–438.62</td>
</tr>
</tbody>
</table>

MCI, mild cognitive impairment; A, additive genetic influences; C, common environmental influences; D, nonadditive/dominance genetic influences; E, nonshared environmental influences; CI, confidence interval; -2LL, -2 log-likelihood; AIC, Akaike’s information criterion. All MCI definitions were based on classification after adjusting for early adult cognitive ability. Model fits were compared with the fit of a saturated model. Df for ACE and ADE models = 1122. Df for AE and CE models = 1123. For all definitions, MCI was treated as a three-level variable: no MCI; single-domain MCI; multiple-domain MCI. Variance components with estimates ≥.10 were not dropped from the models (i.e. set to zero) even if they had 95% confidence intervals that included 0. Models not shown for MCI definitions for which meaningful polychoric correlations could not be computed (see Table 4). Best-fitting models are shown in bold.
The proportion of e4 carriers did not differ between MCI and non-MCI groups (results not shown).

Discussion
In two extensive reviews of MCI,12,13 only a single study included adults under 60 years of age. A 10-item screening test in that study may likely have had ceiling effects in middle-aged adults.17 Our results strongly suggest that MCI can be identified in adults in their 50s provided that tests cover multiple domains and do not have ceiling effects.

Prevalence rates
In an extensive review of older samples, prevalence rates were highly varied (0.5–42%).13 The range in our sample (1.24–64.74%) may be slightly larger because we: (i) examined a larger number of criteria sets \( n = 10 \) from extremely liberal to extremely conservative; (ii) we assessed more cognitive domains

<table>
<thead>
<tr>
<th>MCI definition</th>
<th>Model</th>
<th>Proportion of variance</th>
<th>Model fit statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A (95% CI)</td>
<td>C (95% CI)</td>
</tr>
<tr>
<td>Any MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>ADE</td>
<td>.17 (.00; .75)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.64 (.48; .77)</td>
<td>–</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>ADE</td>
<td>.45 (.00; .70)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.50 (.25; .70)</td>
<td>–</td>
</tr>
<tr>
<td>Typical</td>
<td>ACE</td>
<td>.33 (.00; .50)</td>
<td>.01 (.00; .38)</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.34 (.18; .50)</td>
<td>–</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.21 (.00; .58)</td>
<td>.22 (.00; .49)</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>ADE</td>
<td>.00 (.00; .45)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.33 (.15; .51)</td>
<td>–</td>
</tr>
<tr>
<td>Amnestic MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>ACE</td>
<td>.49 (.00; .87)</td>
<td>.18 (.00; .75)</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Typical</td>
<td>ACE</td>
<td>.22 (.00; .59)</td>
<td>.20 (.00; .50)</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.47 (.00; .69)</td>
<td>.08 (.00; .51)</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.56 (.40; .69)</td>
<td>–</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>ACE</td>
<td>.33 (.00; .60)</td>
<td>.03 (.00; .47)</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.37 (.10; .60)</td>
<td>–</td>
</tr>
<tr>
<td>Non-amnestic MCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite 5</td>
<td>ADE</td>
<td>.20 (.00; .73)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.59 (.41; .74)</td>
<td>–</td>
</tr>
<tr>
<td>Composite 2.5</td>
<td>ADE</td>
<td>.00 (.00; .68)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>.47 (.19; .70)</td>
<td>–</td>
</tr>
<tr>
<td>Typical</td>
<td>ACE</td>
<td>.00 (.00; .31)</td>
<td>.33 (.06; .45)</td>
</tr>
<tr>
<td></td>
<td>CE</td>
<td>–</td>
<td>.33 (.20; .45)</td>
</tr>
<tr>
<td>Comprehensive criteria</td>
<td>ACE</td>
<td>.00 (.00; .42)</td>
<td>.29 (.00; .42)</td>
</tr>
<tr>
<td></td>
<td>CE</td>
<td>–</td>
<td>.29 (.15; .42)</td>
</tr>
<tr>
<td>Conservative criteria</td>
<td>ACE</td>
<td>.00 (.00; .37)</td>
<td>.12 (.00; .32)</td>
</tr>
<tr>
<td></td>
<td>CE</td>
<td>–</td>
<td>.12 (.00; .32)</td>
</tr>
</tbody>
</table>

MCI, mild cognitive impairment; A, additive genetic influences; C, common environmental influences; D, nonadditive/dominance genetic influences; E, nonshared environmental influences; CI, confidence interval; -2LL, -2 log-likelihood; AIC, Akaike’s information criterion. All MCI definitions were based on classification after adjusting for early adult cognitive ability. Model fits were compared with the fit of a saturated model. \( D_f \) for ACE and ADE models = 1123. \( D_f \) for AE and CE models = 1124. For all definitions, MCI was treated as a binary variable: no MCI; MCI (single- or multiple-domain). Variance components with estimates \( \geq .10 \) were not dropped from the models (i.e. set to zero) even if they had 95% confidence intervals that included 0. Model for A-MCI Composite 2.5 not shown because meaningful tetrachoric correlations could not be computed (see Table 4). Best-fitting models are shown in bold.
Prevalence rates for those MCI definitions that ranged from 21.40% to 64.74% are probably overestimates because it is unlikely that over one-fifth of adults in their 50s have MCI. Several multiple-domain definitions had low rates (<10%) for both A-MCI and NA-MCI: 1.33–6.04% for the two composite scores and the Conservative criteria. For single-domain definitions, rates were lower for A-MCI than NA-MCI. These multiple-domain and A-MCI groups may be more likely to be valid because they are less likely to revert to normal on follow-up than other types. In this age range, either of the composite site scores and the Conservative criteria. For single-domain definitions, rates were lower for A-MCI than NA-MCI. These multiple-domain and A-MCI groups may be more likely to be valid because they are less likely to revert to normal on follow-up than other types. 

Comparing number-of-measures-impaired definitions, Jak et al. found in older adults that the Comprehensive criteria were most stable and most likely to be valid. Because base rates and extent of overall impairment will differ, the most successful diagnostic approach may vary in different age cohorts. Even classifications that resulted in relatively high rates may be quite useful. Our classifications may be best thought of as pre-MCI, but that subgroup is still critically important for understanding earlier stages of cognitive aging trajectories and earlier identification of individuals at risk for cognitive decline and dementia.

### Relationship of MCI to early adult GCA

Even after classifying scores based on normative expectations, it may still be difficult to differentiate MCI from longstanding low levels of function. Adjusting for age 20 GCA increases precision; for example, in the entire VETS A, 495 participants had exactly 12 years of education, yet they had a normal distribution with the full range of AFQT scores. At least two processes appear to be at work. Lower age 20 GCA in MCI participants is consistent with early adult GCA as a risk/protective factor decades before for development of MCI. Poorer cognitive functioning—even after adjusting for already lower age 20 GCA—suggests an additional, later-occurring process resulting in trajectories of greater declines compared with non-MCI participants. These results suggest a cognitive reserve phenomenon. Cognitive reserve research has focused largely on much older adults or dementia, but these findings support the need to extend reserve research to healthy adults.

### Genetic influences on MCI

To our knowledge, we have the first evidence that MCI is heritable. Our ability to detect heritability may be due to the combination of adequate sample size plus an extensive test battery. Results were relatively consistent for any-MCI and A-MCI. Average heritabilities based on best-fitting models for any-MCI were .48 (three-level) and .40 (binary). Average heritability that could be estimated for binary A-MCI definitions was .41, but the confidence intervals for two of those four estimates included zero. NA-MCI binary definitions were heritable for composite score, but not number-of-measures-impaired, definitions. More mixed findings for NA-MCI might be due to its being more heterogeneous or perhaps less related to AD than A-MCI. Low prevalence rates may be one reason for some heritabilities having wide confidence intervals that included zero despite moderate-sized estimates. Thus, rather large samples are likely needed for reliable heritability estimates for those MCI definitions. Heritabilities were slightly lower for MCI than AD, but heritabilities for AD also tend to be lower in younger samples. Our findings suggest that any-MCI, and perhaps A-MCI, are appropriate phenotypes for genetic association studies.

In contrast to a study of much older adults, we found no e4-MCI association. However, APOE findings in relatively young, non-demented adults are mixed, and normal-performing e4 carriers in younger samples may well end up with higher rates of conversion compared with non-carriers. Also, these MCI classifications are probably not all ‘MCI due to AD’.

### Strengths and limitations, conclusions

Strengths of the study are MCI assessed at a relatively young age in a national, community-dwelling sample, a comprehensive neuropsychological test battery assessing cognitive domains not always included in prior studies, multiple measures in each domain and increased sensitivity by avoiding ceiling effects in a middle-aged, non-patient sample. We also used multiple MCI definitions. Finally, the fact that results were adjusted for age 20 GCA—providing an objective index of change over time—increases confidence that the MCI classifications do not simply reflect lifelong low overall cognitive ability.

There are also limitations. Generalizability to women or ethnic minorities may be limited. Organization and content of the domains are consistent with many studies in the neuropsychological literature, but neuropsychological domains are not ‘pure’ constructs and tests within a given domain do not necessarily imply homogeneous functional neuroanatomy. Analysis of individual tests or scores might, therefore, provide useful additional information, although those results might be offset by the likely lower reliability of individual scores relative to domain scores.

Although we conducted extensive assessments, there were no formal clinical/diagnostic evaluations. However, participants living independently and flying to study sites demonstrates preserved independence of functional abilities. Results might have
differed if subjective cognitive concerns or informant reports were included, although the utility of such reports has been questioned. According to Petersen, the subjective concern criterion is ‘soft’, ‘but without prior cognitive testing, it is critical for the purpose of excluding individuals with lifelong static cognitive deficits’. In other words, subjective concern is a proxy for prior cognitive data; with prior cognitive data, we had a quantitative index of change from premorbid functioning. Informant reports may also be impractical in community-based studies, but community-based studies are an important complement to selected populations in clinic-based studies.

Whether our MCI classifications are best thought of as MCI or pre-MCI remains to be determined. As with older adults, some will revert to normal and may thus be false positives. Indeed, MCI and MCI-like conditions are likely to represent different phases of ageing with different trajectories and outcomes rather than a single syndrome. Nevertheless, establishing guidelines with community-based samples is particularly important for early identification because few cases in this age range are likely to present in clinical practice.

Although substantial variability in prevalence rates in our study and those of older adults suggests that consensus on MCI criteria is still a way off, evidence that neuropsychologically-defined MCI or pre-MCI can be identified in relatively young, community-dwelling adults who are only in their 50s is encouraging. The impact of this earlier identification could be substantial as it is widely agreed upon that earlier intervention is a key to more effective treatment of AD. Moreover, a 5-year delay of the dementia phase of AD has been estimated to reduce the number of cases by over 50%. Ongoing VETSA follow-ups will provide an opportunity to evaluate the validity of classifications. Finally, we presented the first data showing heritability of MCI. It will be of interest to determine which definitions share the most genetic variance with AD.

**Supplementary Data**
Supplementary data are available at *IJE* online.

**Funding**
This work was supported by the National Institute on Aging at the National Institutes of Health [R01 AG018386, AG022381 and AG022982 to W.S.K.; and AG018384 to M.J.L.], and the Academy of Finland [grant 257075 to E.V.]. It was also, in part, the result of work supported with resources of the VA San Diego Center of Excellence for Stress and Mental Health. The Cooperative Studies Program of the Office of Research & Development of the United States Department of Veterans Affairs has provided financial support for the development and maintenance of the Vietnam Era Twin (VET) Registry. All statements, opinions or views are solely of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government, the National Institute on Aging or the National Institutes of Health.

**Acknowledgments**
The authors gratefully acknowledge the continued cooperation and participation of the members of the VET Registry and their families. Without their contribution, this research would not have been possible. We also appreciate the time and energy of many staff and students on the VETSA projects.

**Conflict of interest:** None declared.

### KEY MESSAGES

- Early identification of MCI is important for reducing or delaying cognitive decline and Alzheimer’s disease; however, participants in most MCI studies are in their 70s, and optimal definitions for impairment cut-offs have yet to be determined.
- After adjusting scores on an extensive test battery for early adult (age 20) general cognitive ability, MCI based on different neuropsychologically-defined criteria could be identified in a community-dwelling sample of middle-aged male twins who were all in their 50s.
- Several MCI definitions were heritable, suggesting that MCI is an appropriate phenotype for genetic association studies.
- It is possible that the identified cases are pre-MCI rather than MCI per se, but such community-based assessment is still important for early identification because it makes it possible to identify at-risk individuals before they present in clinics.
References

The concept of mild cognitive impairment (MCI) was developed with the best of intentions. The term refers to a set of clinically recognizable characteristics that indicate poorer cognitive function than would be expected based on the age and education of the patient, but which are not severe enough to warrant a diagnosis of dementia. The concept is useful, at least in theory, because dementia is inherently a degenerative syndrome. MCI begins mildly but gets steadily, or even increasingly, more severe, and we have few if any effective treatments for most cases. Catching its emergence early or identifying those most vulnerable to MCI would help both in identifying its causes and courses and thus treatments, and in managing cases even before we have effective treatments. The concept of MCI was developed to do just those things.

Commentary: The best-laid plans: the problems and pitfalls of assessing mild cognitive impairment

Wendy Johnson

Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. E-mail: wendy.johnson@ed.ac.uk