Immediate verbal recall and familial dementia risk: population-based study of over 4000 twins

Lindgren, Noora

2019-01

Lindgren , N , Kaprio , J , Rinne , J & Vuoksimaa , E 2019 , ' Immediate verbal recall and familial dementia risk: population-based study of over 4000 twins ' , Journal of Neurology, Neurosurgery and Psychiatry , vol. 90 , no. 1 , pp. 90-97 . https://doi.org/10.1136/jnnp-2018-319122

http://hdl.handle.net/10138/311502
https://doi.org/10.1136/jnnp-2018-319122

Immediate verbal recall and familial dementia risk: population-based study of over 4000 twins

Authors: Noora Lindgren M.Sc.[1,2], Jaakko Kaprio M.D. Ph.D.[3,4], Juha O. Rinne M.D. Ph.D.[1,5], Eero Vuoksimaa Ph.D.[3]

Affiliations: [1] Turku PET Centre, Turku University Hospital and University of Turku, Turku, Finland
[2] Drug Research Doctoral Program, University of Turku, Turku, Finland
[3] Institute for Molecular Medicine Finland, University of Helsinki, Finland
[4] Dept of Public Health, University of Helsinki, Helsinki, Finland
[5] Division of Clinical Neurosciences, Turku University Hospital, Turku, Finland

Word count abstract: 223
Word count paper: 3391

Number of references: 34
Number of tables: 3
Number of figures: 3
Title character count: 93

Corresponding Author:
Noora Lindgren
Turku PET Centre, University of Turku
c/o Turku University Central Hospital
P.O. Box 52 (Kiinamyllynkatu 4-8)
FIN-20521 Turku, FINLAND
Phone +358-445621586
nhsalm@utu.fi

Jaakko Kaprio jaakko.kaprio@helsinki.fi
Juha O. Rinne juha.rinne@tyks.fi
Eero Vuoksimaa eero.vuoksimaa@helsinki.fi

ABSTRACT

Objective

To investigate the effect of familial risk for dementia on verbal learning by comparing cognitively healthy twins who had demented co-twins with cognitively healthy twins who had cognitively healthy co-twins.

Methods

4,367 twins aged ≥65 years including 1,375 twin pairs (533 monozygotic (MZ), 823 dizygotic (DZ) and 19 unknown zygosity pairs) from a population-based Finnish Twin Cohort participated in a telephone assessment for dementia and in a single free-recall trial of a 10-item word list.

Results

Cognitively healthy twins with demented co-twins (*n* = 101 pairs) recalled less words than cognitively healthy twins with cognitively healthy co-twins (*n* = 770 pairs) after adjusting for age, sex, and education, *B* = -0.44, 95% CI [-0.73, -0.14], *P* = .003. The effect size was similar in MZ (*n* = 31) twins (3.88 vs 4.29 words, B = -0.41, 95% CI [-0.96, 0.13]) and DZ (*n* = 66) twins (3.70 vs 4.17 words, B = -0.47, 95% CI [-0.84, -0.10]). The heritability estimate of immediate recall (IR) was 0.37, 95% CI [0.21, 0.43].

Conclusions

The results demonstrate that familial risk for dementia is reflected in the IR performance of cognitively healthy older persons. The finding of poorer IR performance in non-affected siblings compared to the general population, together with substantial

heritability of IR, supports IR as a useful endophenotype for molecular genetic studies of dementia.

Keywords: Dementia, Memory and Learning Tests, Verbal Learning, Episodic Memory, Twin Study

INTRODUCTION

Impaired verbal learning and memory performance, which is commonly measured with word-list learning tests, is often the earliest clinical sign of Alzheimer's disease (AD). Immediate and delayed recall of word lists have been identified as early predictors of all-cause or AD dementia.[1–5]. Previous research suggests that immediate recall (IR) scores may be the most sensitive indicator of risk in the earliest preclinical AD stages and delayed recall scores may be a better indicators in later stages.[1,6] Analysis of the serial position effect in immediate word-list recall may provide added sensitivity in identifying asymptomatic persons at risk for AD.[7] The serial position effect refers to the tendency to remember more items in the beginning (primacy) and end (recency) of a list than in the middle, thus producing the characteristic U-shaped serial position curve.

Genetic factors have a significant role in verbal learning and memory and in age-related dementias, with particularly AD having a substantial heritability. APOE ε4 allele is the strongest genetic risk factor for decline in immediate and delayed verbal recall and AD, but other genes also affect the risk.[8,9] In addition to heritability estimates, studies of monozygotic (MZ) and dizygotic (DZ) twins can investigate the effect of familial risk of dementia on cognitive performance by controlling for unmeasured shared genetic (in MZ's) and environmental (in DZ's and MZ's) influences. We studied a population-based sample of older Finnish twins who participated in a telephone interview of cognitive status including a word-list learning measure. We examined if the IR performance of

cognitively healthy twins with demented co-twins differs from cognitively healthy twins with cognitively healthy co-twins. We also estimated the heritability of immediate free recall of a single trial 10-item word list.

METHODS

Participants and study design

The participants for telephone screening of cognition were recruited from the older Finnish Twin Cohort (FTC) study, consisting of same-sex twin pairs born prior to 1958. The FTC database is described in detail earlier.[10] In 1975 and 1981, the twins participated in a postal questionnaire data collection with high response rates of 89% and 84%, respectively. In 1999 - 2007, FTC twins aged ≥65 years and born before 1938 were asked to participate in a telephone interview. The second wave of interviews was carried out during 2013 - 2017 for FTC twins aged ≥65 years and born in 1938 - 1944. The overall participation rate was 67% (4,403/6,572). Among the cohort of 6,572 twins, DNA-based zygosity was available for 1,542 twins, for others zygosity was determined with a validated questionnaire.[11] Three groups of twins were created from the population-based sample: cognitively healthy twins with demented co-twins, demented twins with cognitively healthy co-twins, and the cognitively healthy twins with cognitively healthy co-twins. A flowchart of the study design with the number of twins included in the analyses is shown in figure 1.

The telephone interview protocol was approved by the Ethics Committee of the Hospital District of Southwest Finland and informed consent was obtained from participants. Questionnaire studies in 1975 and 1981 were approved by the National Board of Health and answering the questionnaire was considered as consent.

Measures

The telephone interview included the telephone assessment for dementia (TELE)[12] and the Telephone Interview for Cognitive Status (TICS)[13]. Their validation in Finnish and detailed description can be found elsewhere [14]. The interviews were carried out by trained personnel with a medical background. Cognitive status was defined based on the TELE score (range 0 - 20.0), not on diagnostic criteria. We used Finnish TELE cut-off scores for healthy cognition (> 17.5) and dementia (< 16.0).[14] As a measure for IR, we used the free recall of a 10-item word list in TICS. The primary measure was the total number of immediately recalled words after a single administration of the list. We also studied the number of words recalled according to the serial position: primacy (first three words) and recency (last three words). Education was based on the 1975 and 1981 self-report questionnaire and included 8 categories. This information was transformed into years of education and used as a continuous variable. In case of differences between the 1975 and 1981 questionnaires, the higher level of education provided was chosen.

Statistical analysis

The statistical analyses were conducted using Stata version 14.2 (Stata Corp., College Station, TX). Two-tailed P-values < .05 indicated statistical significance. The primary research question was whether IR performance of cognitively healthy twins who had demented co-twins is different compared to cognitively healthy twins who had cognitively healthy co-twins. In the between family-analyses, we used linear regression to compare the total number of recalled words between the groups. In the within-family analyses, a linear conditional fixed effects regression model was used to examine differences in total scores. Ordered logistic regression was used for the analysis of the number of recalled primacy and recency words. All regression analyses were controlled for age, sex, and years of education. The twin structure of the data was considered in all

analyses by using robust standard errors adjusted for family relatedness.[15] In linear regression analyses, adjusted means and unstandardized coefficients with 95% CIs were reported. Assumptions of regression analyses were verified.

MZ and DZ male and female twin correlations for IR performance were calculated using age-adjusted standardized residuals of total scores. Twin correlations were calculated with age- and sex-adjusted measures when including both men and women. In the twin design, it is possible to decompose the phenotypic variance into additive genetic (A), common environmental (C) and unique environmental (E) variances. The ACE model is based on the assumption that A effects correlate 1.0 in MZ twins because they are genetically identical at the sequence level, whereas the correlation is 0.5 in DZ twins who share (like siblings), on average, 50% of their segregating genes. C effects refer to all environmental influences that make twins within a pair similar and correlate 1.0 both in MZ and DZ pairs. E effects refer to all environmental influences that make twins within a pair different from each other and are thus uncorrelated. E effects also include measurement errors. We used a maximum likelihood based structural equation modeling approach to estimate the relative proportion of A (heritability), C and E effects.[16] We adjusted IR measure for age and sex and scores were standardized with a mean of 0 and SD of 1 before model fitting.

RESULTS

IR performance in the whole population-based cohort

The 4,367 participants with complete data had a mean age of 74.1 years (*SD* = 4.1, range 65 - 97), mean education of 7.9 years (*SD* = 2.9, range: 5 - 16), and 48.6% were women. The participants were more often men (51% vs 43%) and slightly more educated (7.9 vs 7.2 years) but of similar age in comparison to non-participants. The

mean TELE score was 18.0 (*SD* = 2.2, interquartile rang (IQR): 17.0 - 19.5). The mean total score for the word list recall was 3.7 words (*SD* = 1.7, IQR: 3 - 5, with a normal distribution). The descriptive statistics of word list recall measures are presented in figure 2. Better word list recall was associated with higher education (Spearman's rho = 0.30, P < .001), lower age (Pearson's r = -0.21, P < .001), and female sex (P < .001). Participants showed the usual U-shaped serial position curve that decreased overall as a function of decreasing cognitive status, decreasing education or increasing age (figure 2).

IR performance in cognitively concordant and discordant twin pairs

Out of the 1,375 pairs, 110 pairs were discordant for dementia, 770 pairs were concordant for healthy cognition, and 39 pairs were concordant for dementia. Table 1 shows the characteristics of twins in pairs discordant for dementia and in pairs concordant for healthy cognition. Twins in discordant pairs were on average slightly older and had less years of education than twins in cognitively healthy pairs.

Table 1 Characteristics of cognitively discordant twin pairs and pairs concordant for healthy cognition

	Cognitively discordant pairs (n = 101)			Cognitively healthy pairs (n = 770)	
	Demented twins (n = 101)	Cognitively healthy twins (n = 101)	P value[#]	Cognitively healthy twins (n = 1540)	P value[##]
Age, years	74.5 (4.4)	74.5 (4.4)	.962	72.3 (2.9)	<.001
Education, years	7.0 (1.9)	7.6 (2.7)	.018	8.7 (3.3)	<.001

Sex [% female]	48.5	48.5		55.2	.191
Zygosity					.010
Monozygotic [%]	31	31		43	
Dizygotic [%]	65	65		56	
Unknown [%]	4	4		1	
TELE	13.2 (2.2)	18.9 (0.7)	<.001	19.1 (0.7)	<.001
Total IR score	2.7 (1.7)	3.5 (1.5)	.001	4.3 (1.7)	.003
Primacy words [median, IQR]	1 (0, 1)	1 (1, 2)	<.001	2 (1, 2)	.561
Recency words [median, IQR)	1 (0, 2)	1 (1, 2)	.192	2 (1, 3)	.157

Abbreviations: TELE, Telephone assessment for dementia. IR, immediate recall.

NOTE. Values are mean (SD) unless otherwise stated. The total IR score is the number of recalled words in a single free-recall trial of a 10-item word list. Cognitive status is defined based on telephone assessment of dementia (TELE) cut-off scores (dementia: TELE < 16; healthy cognition TELE > 17.5).

Statistical significance of the difference in age, education, and TELE between co-twins tested with paired t test. The difference in IR between co-twins tested with linear conditional fixed-effects regression.

Statistical significance of the difference between the cognitively healthy twin groups with two-sample t-test. Pearson's chi-squared test used for categorical variables.
The differences in IR between the cognitive healthy groups with linear regression and differences in primacy and recency words with ordered logistic regression, adjusted for age, education, and family relatedness.

Figure 3 and table 2 summarize the results of the differences in total scores between the three groups: cognitively healthy twins with demented co-twins, demented co-twins, and cognitively healthy twins with cognitively healthy co-twins. Cognitively healthy twins from discordant pairs had better IR than their demented co-twins, respectively 3.79 vs 3.05 words, *B* = 0.74, 95% CI [0.29, 1.19], *P* = .001. Cognitively healthy twins with demented co-twins had poorer total scores than cognitively healthy twins with healthy co-twins after adjusting for age, sex, and education, respectively 3.79 vs 4.23 words, *B* = -0.44, 95% CI [-0.73, -0.14], *P* = .003. This difference with a Cohen's *d* of 0.3 corresponded to a difference of about 6 years of age on IR performance or the difference between men and women. The difference in IR performance remained statistically significant after further adjusting for TELE score (*P* = .008).

We did not detect a statistically significant difference specifically in primacy (*B* = -0.11, 95% CI [-0.50, 0.27], *P* = .561) or recency words (*B* = -0.25, 95% CI [-0.59, 0.10], *P* = .157) between cognitively healthy twins with demented co-twins and cognitively healthy twins with healthy co-twins.

Differences in total scores between cognitively healthy twins from cognitively discordant pairs and concordant pairs were also analyzed separately in MZ and DZ twins (31 discordant and 328 concordant MZ pairs, 66 discordant and 432 concordant DZ pairs). The differences were of similar magnitude in both MZ and DZ twins (MZ: 3.88 vs 4.29, *B* = -0.41, 95% CI [-0.96, 0.13], *P* = .134; DZ: 3.70 vs 4.17, *B* = -0.47, 95% CI [-0.84, -0.10], *P* = .012), but the coefficient was not statistically significant in MZ twins.

Table 2 Between-family analysis results of the differences in total scores between twins in cognitively discordant pairs and pairs concordant for healthy cognition

	All twins[a]			MZs[b]			DZs[c]		
	B	95% CI	P	B	95% CI	P	B	95% CI	P
Group									
Cognitively healthy twins with cognitively healthy co-twins	Reference			Reference			Reference		
Cognitively healthy twins with demented co-twins	-0.44	-0.73, -0.14	.003	-0.41	-0.96, 0.13	.134	-0.47	-0.84, -0.10	.012
Demented co-twins	-1.18	-1.51, -0.84	<.001	-1.17	-1.78, -0.55	<.001	-1.16	-1.58, -0.74	<.001
Sex									
Male	Reference			Reference			Reference		
Female	0.49	0.33, 0.65	<.001	0.43	0.17, 0.68	.001	0.52	0.31, 0.73	<.001
Age	-0.07	-0.10, -0.05	<.001	-0.08	-0.11, -0.04	<.001	-0.07	-0.11, -0.03	<.001
School years	0.13	0.10, 0.16	<.001	0.15	0.10, 0.19	<.001	0.12	0.09, 0.15	<.001

Abbreviations: B, unstandardized coefficient; MZ, monozygotic; DZ, dizygotic.

[a] 101 discordant pairs and 770 pairs concordant for healthy cognition included.

[b] Only MZs included (31 discordant and 328 pairs concordant for healthy cognition).

[c] Only DZs included (66 discordant and 432 pairs concordant for healthy cognition).

Heritability of single free-recall trial of a 10-item word list

Within-twin pair correlations of IR performance were higher in MZ twins compared to DZ twins (rMZ = 0.38, rDZ = 0.18). Correlations were similar when analyzing men and women separately (table 3). The univariate biometrical model yielded a heritability estimate of 0.37, 95% CI [0.21, 0.43]. C effects were estimated at 0, 95% CI [0, 0.12] and E effects were 0.63, 95% CI [0.57, 0.69].

Table 3 Within twin pair correlations of immediate recall (IR) performance in 533 monozygotic (MZ) and 823 dizygotic (DZ) pairs

	MZ				DZ			
	Men (*n* = 250)		Women (*n* = 283)		Men (*n* = 411)		Women (*n* = 412)	
	r	95% CI	r	95% CI	r	95% CI	r	95% CI
IR score	0.39	0.28, 0.49	0.37	0.27, 0.47	0.19	0.09, 0.28	0.17	0.08, 0.26

The number of pairs is shown in brackets. The IR score is the number of recalled words in a single free recall trial of a 10-word list.

DISCUSSION

We used a telephone interview to evaluate cognitive status and immediate free recall performance in a population-based sample of older twins that was highly representative of the older Finnish population. We first examined the recall performance of a 10-item word list and its association with demographic factors in the whole cohort. As expected, a lower IR performance was associated with lower cognitive status, older age, lower education, and male sex.

Our main interest was to examine whether increased risk for dementia would be reflected in the IR performance of cognitively healthy older twins who had a demented twin sibling. Twin studies have indicated that genetic effects account for about half of the variance for any dementia and even more for AD.[17] We also show genetic effects on IR performance in the present sample. In addition to MZ twin pairs sharing all their genes and DZ twins sharing on average 50% of their segregating genes, twins growing up in the same family share many environmental risk factors that may affect cognitive ability even late in life.

The primary result was that cognitively healthy twins, who had a demented co-twin, showed poorer IR performance than healthy comparison subjects who were cognitively healthy twins, whose twin brother or sister was not demented. The analysis was controlled for age, sex, and education. The difference in total recall was subtle but statistically significant with a modest effect size. Our results indicate that poorer IR performance may be a leading indicator for increased risk for dementia in the presence of cognitively normal performance as indicated by a more general screening instrument. The 10-item word list was presented only once over the telephone and the test is difficult even for cognitively healthy older individuals. The commonly used word list learning tests, 16-item California Verbal Learning Test (CVLT), 15-item Rey Auditory Verbal Learning Test (AVLT), 10-item word list recall tests from CERAD and ADAS-Cog, measure IR performance as the total number of recalled words over three or five immediate free recall trials. It is possible that a larger effect size would have been detected in our study by using multiple trials of the 10-item word list.

The poorer IR performance in the cognitively healthy twins with demented co-twins was associated with familial (genetic and shared environmental) risk for dementia. We did not detect evidence that cognitively healthy MZ twins from discordant pairs would have poorer IR performance in relation to cognitively healthy DZ twins from discordant pairs. The effect

was similar in direction and in magnitude in analyses of both MZ and DZ twins. The sample size in MZ pairs was less than half of that in DZ twin pairs, and a larger number of discordant MZ pairs is needed for more definite conclusions.

The cognitively discordant twin pairs may be underrepresented in the study sample because cognitively impaired individuals, especially the most severely impaired, may be more often not willing or able to participate in the study compared to cognitive healthy individuals. This may potentially bias the estimates of the effect. We assume that smaller number of identified cases would more likely, if anything, decrease our possibilities to detect a poorer recall performance in the cognitively healthy co-twins from discordant pairs.

Serial position effects in word list recall may possibly be used to differentiate normal aging from pathological aging. Diminished primacy effect (i.e. remembering less words from the beginning of a word list) has been observed in AD and MCI and shown to improve discrimination between progressive and non-progressive MCI.[18] It is possible that serial position effects may also predict risk for dementia in cognitively healthy individuals. La Rue et al.[7] reported that asymptomatic middle-aged persons with a family history of AD have poorer primacy effect but no difference in the total immediate or delayed recall scores of AVLT, in comparison to individuals without a family history of AD. In our study, we did detect a poorer total immediate recall score but no significant difference specifically in the primacy performance in the cognitively healthy twins with demented co-twins compared to cognitively healthy twins with non-demented co-twins. The major difference between our study and the study by La Rue et al. is that we examined the IR performance of older people with increased familial risk for dementia not restricted to AD. In addition, we used a single-trial test and a shorter word list (10-item vs 15-item) that might be less sensitive for differences in recall according to serial position regions.

More than a third of the variance, 37%, of our IR measure was explained by genetic effects. This is in line with an earlier study reporting a heritability of 0.36 for word-list learning measure including multiple trials (CVLT) .[19] However, the first trial of the 16-item word list (CVLT) has been reported to show zero heritability indicating that a single-trial administration of a word list may not be reliable enough to show consistent heritability.[20] In our study, the 10-item word list was presented only once but still showed substantial heritability. The single trial of a shorter word list may be more reliable compared to a single presentation of a longer supraspan list (i.e., exceeding typical working memory capacity).

Our results support poorer word list learning as an early marker of dementia risk. The weaker performance may reflect a vulnerability trait, a certain cognitive phenotype, that may increase the risk of developing dementia. It is also possible that the subtle alteration in IR is causally linked to the occurrence of underlying progressive brain pathologies, and thus may represent a preclinical sign of pathology. However, to make this inference further studies with longitudinal follow-up and information also on pathological markers are needed. Regional atrophy in lateral temporoparietal, posterior cingulate and precuneus, and frontal regions has been associated with the performance of the first immediate recall trial of AVLT and ADAS-Cog in mild AD patients.[21] Several of these brain regions, especially the precuneus and posterior cingulate, are associated with the default mode network and are affected very early during the AD process by high amyloid-beta burden, [22,23] and disrupted functional connectivity.[24]

The major limitations of our study are related to the definition of dementia. We studied the effect of all-cause dementia risk on IR performance. The study design and assessment measures limited us from examining the effect by dementia subtypes. Most of the dementia cases are presumably due to AD.[25] AD has a long preclinical period and pathophysiological changes in the brain can begin even decades before the disease is

diagnosed.[26] Another important cause for cognitive impairment and dementia in the elderly is cerebrovascular disease including vascular dementia (VaD). VaD may also have a long preclinical period during which immediate free recall performance declines.[27] Vascular risk factors are major risk factors for both AD and VaD. AD-related and cerebrovascular pathologies also show considerable overlap. In fact, for elderly individuals, the presence of a single disease may rather be the exception because AD- and non-AD type pathologies coexist very frequently.[25,28] Hence, the assessment of dementia risk for any reason, not only disease-specific, in older adults may be valuable.

The definition of dementia was done based on a validated instrument TELE that has adequate properties for the screening of dementia.[12,14] TELE scores have shown high correlation with the Mini-Mental State Examination and Clinical Dementia Rating Scale Sum of Boxes scores.[14] The telephone-based definition of dementia also identifies well those who report having difficulties in daily activities.[29] Nevertheless, a major limitation of our study is that the definition of dementia is not based on clinical diagnostic criteria but on the telephone assessment. False positive and negative findings are a concern when using any brief cognitive screening instrument and cut-off scores to detect dementia. The TELE cut-off score for dementia (< 16.0) provides the specificity of 100% and sensitivity of 77% for dementia based on an earlier case-control study.[14] In a population-based study, the cut-off score for dementia (<16.0) showed a lower specificity of 90%.[12] Consequently, it is likely that the demented twins in discordant pairs include false positive participants. Most false positive cases are likely cases that suffer from cognitive impairment and who subsequently develop dementia, but in some cases a low test score may be due to a mental disorder, impaired hearing, low education or low intelligence.[12] The possible inclusion of false positive cases may dilute the difference between cognitively healthy twins with demented co-twins and cognitively healthy twins with non-demented co-twins.

On the other hand, it is also possible that the risk of a false-positive result might be associated with common familial factors. The cut-off score for healthy cognition (> 17.5) has the sensitivity of 97% and specificity of 69%.[11] Thus, the twins who are defined as cognitively healthy in discordant and concordant pairs are likely to include very few demented individuals.

The participants were inquired about their hearing problems and were asked to confirm that they were in a quiet place and had no external memory aids at the beginning of the interview. Nevertheless, the telephone interview is limited by the restricted control over external distractors. The advantages of the telephone-based assessment of cognition included low cost and easy accessibility. The overall strengths of our study are the large population-based sample and the unique dementia-discordant twin pair design.

Our results support the use of IR as an endophenotype. Endophenotypes are defined as quantitative, subclinical traits that are associated with the illness, are heritable, are state independent, co-segregate with the illness within families, and are found in non-affected family members at a higher rate than in the general population.[30] IR impairment has a well-known association with dementia, especially with AD.[31,32] Our results support the finding that poorer IR performance is found at a higher rate in non-affected family members compared to the general population. An increasing number of genome-wide association studies are employing AD-related endophenotypes, such as cerebrospinal fluid, imaging, and neuropsychological traits.[33,34] Identifying genetic variants that are associated both with IR endophenotype and increased risk of age-related dementias such as AD may provide a useful approach to disentangle the mechanisms of risk genes.[34] Our results show that even a simple word list learning task administered by telephone may be useful in such studies.

In conclusion, we found that familial risk of cognitive impairment is reflected in the IR performance of older persons who are cognitively normal based on a general cognitive measure. Our study also demonstrates the possible utility of a telephone-administered word list learning task.

Acknowledgements

We gratefully acknowledge the assistance of the following persons: Kristiina Saanakorpi and Ulla Kulmala-Gråhn for interviewing the participants, Pia Ruokolinna for data entry, and Kauko Heikkilä for data management. We are grateful for all the twins who participated in the Finnish Twin Cohort study.

Author statement

N. Lindgren: drafting the manuscript, study concept and design, analysis and interpretation of data, acquisition of data, statistical analysis. J. Kaprio: recruitment of the study cohort, supervision of data collection, interpretation of data. JO. Rinne: study concept and design, interpretation of data, revising the manuscript, study supervision, obtaining funding. E. Vuoksimaa: drafting/revising the manuscript, study concept and design, analysis and interpretation of data, statistical analysis, study supervision. All authors have provided critical comments on the manuscript and approved the final version.

Funding

This work was supported by Sigrid Juselius Foundation and Finnish Governmental Research Funding. N. Lindgren was supported by the Finnish Cultural Foundation, Yrjö

Jahnsson Foundation, Turku University Foundation, and Finnish Brain Foundation. J. Kaprio was supported by the Academy of Finland (grants 265240 & 263278). E. Vuoksimaa was supported by Juho Vainio Foundation. The funders had no role in study design, data collection, data analysis, data interpretation, or in writing of the report.

Author disclosures

N. Lindgren, J. Kaprio and E. Vuoksimaa report no disclosures. JO. Rinne serves as a neurology consultant for Clinical Research Services Turku (CSRT Oy).

Data sharing

Due to the consent given by study participants and the high degree of identifiability, data cannot be made publicly available. Data are available through the Institute for Molecular Medicine Finland (FIMM) Data Access Committee (DAC) for authorized researchers who have IRB/ethics approval and an institutionally approved study plan. For more details, please contact the FIMM DAC (fimm-dac@helsinki.fi).

REFERENCES

1 Bilgel M, An Y, Lang A, *et al.* Trajectories of Alzheimer disease-related cognitive measures in a longitudinal sample. *Alzheimer's Dement* 2014;**10**:735–42. doi:10.1016/j.jalz.2014.04.520

2 Blacker D, Lee H, Muzikansky A, *et al.* Neuropsychological measures in normal individuals that predict subsequent cognitive decline. *Arch Neurol* Published Online First: 2007. doi:10.1001/archneur.64.6.862

3 Landau SM, Harvey D, Madison CM, *et al.* Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology* 2010;**75**:230–8. doi:10.1212/WNL.0b013e3181e8e8b8

4 Tierney MC, Yao C, Kiss A, *et al.* Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology* 2005;**64**:1853–9. doi:10.1212/01.WNL.0000163773.21794.0B

5 Tierney MC, Moineddin R, McDowell I. Prediction of all-cause dementia using neuropsychological tests within 10 and 5 years of diagnosis in a community-based sample. *J Alzheimer's Dis* 2010;**22**:1231–40. doi:10.3233/JAD-2010-100516

6 Rabin LA, Paré N, Saykin AJ, *et al.* Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Aging, Neuropsychol Cogn* 2009;**16**:357–76. doi:10.1080/13825580902825220

7 Rue A La, Hermann B, Jones JE, *et al.* Effect of parental family history of Alzheimer's disease on serial position profiles. *Alzheimer's Dement* 2008;**4**:285–90. doi:10.1016/j.jalz.2008.03.009

8 Desikan RS, Fan CC, Wang Y, *et al.* Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS Med* 2017;**14**. doi:10.1371/journal.pmed.1002258

9 Andrews SJ, Das D, Cherbuin N, *et al.* Association of genetic risk factors with cognitive decline: The PATH through life project. *Neurobiol Aging* 2016;**41**:150–8. doi:10.1016/j.neurobiolaging.2016.02.016

10 Kaprio J, Koskenvuo M. Genetic and environmental factors in complex diseases:

The older Finnish Twin Cohort. *Twin Res* 2002;**5**:358–65. doi:10.1375/twin.5.5.358

11 Sarna S, Kaprio J, Sistonen P, *et al.* Diagnosis of twin zygosity by mailed questionnaire. *Hum Hered* 1978;**28**:241–54. doi:10.1159/000152964

12 Gatz M, Reynolds CA, John R, *et al.* Telephone screening to identify potential dementia cases in a population-based sample of older adults. *Int Psychogeriatrics* 2002;**14**:273–89. doi:10.1017/S1041610202008475

13 Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. *Neuropsychiatry Neuropsychol Behav Neurol* 1988;**1**:111–7.http://journals.lww.com/cogbehavneurol/Abstract/1988/00120/The_Telephone_Interview_for_Cognitive_Status.4.aspx

14 Järvenpää T, Rinne JO, Räihä I, *et al.* Characteristics of two telephone screens for cognitive impairment. *Dement Geriatr Cogn Disord* 2002;**13**:149–55. doi:10.1159/000048646

15 Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics. 2000;**56**:645–6. doi:10.1111/j.0006-341X.2000.00645.x

16 Neale M, Cardon L. *Methodology for genetic studies of twins and families*. 1st ed. Dordrecht: : Kluwer 1992.

17 Gatz M, Reynolds CA, Fratiglioni L, *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 2006;**63**:168–74. doi:10.1001/archpsyc.63.2.168

18 Cunha C, Guerreiro M, de Mendonca A, *et al.* Serial position effects in Alzheimer's disease, mild cognitive impairment, and normal aging: predictive value for

conversion to dementia. *J Clin Exp Neuropsychol* 2012;**34**:841–52. doi:10.1080/13803395.2012.689814

19 Panizzon MS, Lyons MJ, Jacobson KC, *et al.* Genetic Architecture of Learning and Delayed Recall: A Twin Study of Episodic Memory. *Neuropsychology* Published Online First: 2011. doi:10.1037/a0022569

20 Kremen WS, Panizzon MS, Franz CE, *et al.* Genetic complexity of episodic memory: A twin approach to studies of aging. *Psychol Aging* 2014;**29**:404–17. doi:10.1037/a0035962

21 Wolk DA, Dickerson BC. Fractionating verbal episodic memory in Alzheimer's disease. *Neuroimage* Published Online First: 2011. doi:10.1016/j.neuroimage.2010.09.005

22 Palmqvist S, Schöll M, Strandberg O, *et al.* Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nat Commun* Published Online First: 2017. doi:10.1038/s41467-017-01150-x

23 Villain N, Chételat G, Grassiot B, *et al.* Regional dynamics of amyloid-β deposition in healthy elderly, mild cognitive impairment and Alzheimer's disease: A voxelwise PiB-PET longitudinal study. *Brain* Published Online First: 2012. doi:10.1093/brain/aws125

24 Hedden T, Van Dijk KRA, Becker JA, *et al.* Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. *J Neurosci* Published Online First: 2009. doi:10.1523/JNEUROSCI.3189-09.2009

25 Neuropathology Group. Medical Research Council Cognitive Function and Aging Study Z, Mirra S, Heyman A, *et al.* Pathological correlates of late-onset dementia in

a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet (London, England)* 2001;**357**:169–75. doi:10.1016/S0140-6736(00)03589-3

26 Villemagne VL, Burnham S, Bourgeat P, *et al.* Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *Lancet Neurol* 2013;**12**:357–67. doi:10.1016/S1474-4422(13)70044-9

27 Laukka EJ, MacDonald SWS, Fratiglioni L, *et al.* Preclinical cognitive trajectories differ for Alzheimer's disease and vascular dementia. *J Int Neuropsychol Soc* 2012;**18**:191–9. doi:10.1017/S1355617711001718

28 Jellinger KA, Attems J. Challenges of multimorbidity of the aging brain: a critical update. J. Neural Transm. 2015;**122**:505–21. doi:10.1007/s00702-014-1288-x

29 Vuoksimaa E, Rinne JO, Lindgren N, *et al.* Middle age self-report risk score predicts cognitive functioning and dementia in 20–40 years. *Alzheimer's Dement Diagnosis, Assess Dis Monit* 2016;**4**:118–25. doi:10.1016/j.dadm.2016.08.003

30 Gottesman II, Gould TD. The endophenotype concept in psychiatry: Etymology and strategic intentions. Am. J. Psychiatry. 2003;**160**:636–45. doi:10.1176/appi.ajp.160.4.636

31 Almkvist O, Fratiglioni L, Agüero-Torres H, *et al.* Cognitive Support at Episodic Encoding and Retrieval: Similar Patterns of Utilization in Community-Based Samples of Alzheimer's Disease and Vascular Dementia Patients. *J Clin Exp Neuropsychol* 1999;**21**:816–30. doi:10.1076/jcen.21.6.816.862

32 Economou A, Routsis C, Papageorgiou SG. Episodic memory in Alzheimer disease, frontotemporal dementia, and dementia with lewy bodies/Parkinson disease dementia: Disentangling retrieval from consolidation. *Alzheimer Dis Assoc Disord* 2016;**30**:47–52. doi:10.1097/WAD.0000000000000089

33 Chung J, Wang X, Maruyama T, *et al.* Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. *Alzheimer's Dement* Published Online First: 2017. doi:10.1016/j.jalz.2017.11.006

34 Barral S, Bird T, Goate A, *et al.* Genotype patterns at PICALM, CR1, BIN1, CLU, and APOE genes are associated with episodic memory. *Neurology* 2012;**78**:1464–71. doi:10.1212/WNL.0b013e3182553c48

Figure 1 Flowchart of the study and number of twins included in the analyses. Cognitive status is defined based on the TELE cut-off scores (dementia: TELE < 16; healthy cognition TELE > 17.5). Abbreviations: *MZ*, monozygotic; *DZ*, dizygotic; *XZ*, unknown zygosity; TELE, telephone assessment for dementia.

Figure 2 Episodic memory (EM) performance according to sex, cognitive status, and education among the 4,367 participants. The bars represent the mean number of immediately recalled words with +/- SD. The darker gray bars represent men, and the lighter gray women. **A** Word list recall according to sex and cognitive status. Cognitive status is defined based on a telephone assessment for dementia (cut-off score for dementia: TELE < 16; healthy cognition TELE > 17.5). **B** Proportion of correctly recalled words according to cognitive status and as a function of serial position. The triangles

represent the serial position curve of those cognitively healthy and the squares of those with dementia. **C** Immediate word list recall according to sex and education categories of 6 years or less, 7-12 years, and 13 or more years of formal education. **D** Proportion of correctly recalled words according to education category and as a function of serial position. The squares represent the serial position curve of participants with 6 years or less of formal education, circles 7-12 years, and triangles 13 or more years.

Figure 3 Episodic memory (EM) performance of twins in cognitively discordant pairs and cognitively healthy pairs. EM is measured with immediate free recall of a 10-word list. **A** Age, sex and education adjusted means of recalled words with 95% CIs for the demented twins, their cognitively healthy co-twins, and cognitively healthy twins with healthy co-twins. **B** Proportion of correctly recalled words as a function of serial position. Closed squares represent the group of demented twins, open circles the group of cognitively healthy twins with demented co-twins, and closed triangles the cognitively healthy twins with healthy co-twins.