GxE interaction influences trajectories of hand grip strength

Inge Petersen  
The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark

Nancy L. Pedersen  
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Taina Rantanen  
Gerontontology Research Center and Department of Health Sciences, University of Jyväskylä, Jyväskylä, Finland

William S. Kremen  
Department of Psychiatry and Center for Behavioral Genomics, University of California, San Diego, La Jolla, CA, USA  
Center of Excellence for Stress and Mental Health, VA San Diego Healthcare Center, La Jolla, CA, USA

Wendy Johnson  
Department of Psychology and Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

Matthew S. Panizzon  
Department of Psychiatry and Center for Behavioral Genomics, University of California, San Diego, La Jolla, CA, USA

Lene Christiansen  
The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark

Carol E. Franz  
Department of Psychiatry and Center for Behavioral Genomics, University of California, San Diego, La Jolla, CA, USA

Matt McGue  
Department of Psychology, University of Minnesota, Minneapolis, MN, USA  
The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark

Kaare Christensen  
The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark  
Department of Clinical Genetics, Odense University Hospital, Odense, Denmark  
Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense, Denmark

Nayla R. Hamdi  
Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA

Robert F. Krueger  
Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA

Chandra Reynolds  
Department of Psychology, University of California – Riverside, Riverside, CA, USA

For the IGEMS consortium
Abstract

Age-related decline in grip strength predicts later life disability, frailty, lower well-being and cognitive change. While grip strength is heritable, genetic influence on change in grip strength has been relatively ignored, with non-shared environmental influence identified as the primary contributor in a single longitudinal study. The extent to which gene-environment interplay, particularly gene-environment interactions, contributes to grip trajectories has yet to be examined. We considered longitudinal grip strength measurements in seven twin studies of aging in the IGEMS consortium. Growth curve parameters were estimated for same-sex pairs, aged 34-99 (N=10,681). Fisher’s test for mixture distribution of within-monozygotic twin-pair differences (N=1,724) was performed on growth curve parameters. We observed significant gene-environment interaction on grip strength trajectories. Finally, we compared the variability of within-pair differences of growth curve parameters by APOE haplotypes. Though not statistically significant, the results suggested that APOE ε2ε2/ε2ε3 haplotypes might buffer environmental influences on grip strength trajectories.

Word count: 149

Keywords: Grip strength, gene-environment interaction, twins, APOE
Introduction

Age-related loss of skeletal muscle mass has been associated with several adverse age-related outcomes including higher risk of mortality (Cruz-Jentoft et al., 2010b). The age-related loss of muscle mass is due to decreasing number and size of myofibres, but the process can be slowed down or even reversed by exercise and dietary supplements (Sayer et al., 2013). Hand grip strength has been shown to correlate with elbow flexion strength as well as knee and trunk extension strength (Tiainen et al., 2004), and it has been recommended as the best technique for measuring muscle strength (Cruz-Jentoft et al., 2010a). However, in a recent study of elderly women, the usefulness of grip strength as a proxy measure of muscle strength in lower extremities was questioned (Felicio et al., 2014). The existing literature provides evidence that grip strength reflects a mixture of genetic predispositions, environmental factors, and diseases. Indeed, grip strength has been suggested to be a more powerful single marker of frailty than chronological age in a group of elderly (Syddall et al., 2003).

Grip strength is easily measured in the clinic or at home-visits and is among the most studied phenotypes in the literature on phenotypes of aging. A comprehensive literature provides evidence that grip strength is a strong predictor of adverse outcomes in elderly people. Poor grip strength has been demonstrated to predict disability in activities of daily living (ADL) (Taekema et al., 2010; den Ouden et al., 2013; Rantanen et al., 1999), persisting depression and anxiety disorders (van Milligen et al., 2012), depression (Gatz et al., 2010), lower cognitive performance (Sternang et al., 2015), reduced social and leisure activities among the oldest old (Taekema et al., 2010), higher risk of being hospitalized (Legrand et al., 2014) and longer stays at hospital (Mendes et al., 2014). Moreover, several studies have established an association between low grip strength and higher mortality rates (Legrand et al., 2014; Rantanen et al., 2012; Cooper et al., 2014; Rantanen et al., 2000; Rantanen et al., 2003).

Grip strength is a measure that captures early and recent exposures and depends on internal factors such as age and sex (Nahhas et al., 2010; Rantanen et al., 2000; Frederiksen et al., 2006; Sternang et al., 2014). Men were significantly stronger, but they also demonstrated steeper decline compared to women (Frederiksen et al., 2006; Nahhas et al., 2010; Sternang et al., 2014). Moreover, Nahhas et al (Nahhas et al., 2010) found that the decline in grip strength begins in midlife and continues throughout life, which is consistent with another study suggesting accelerating declines in late life for men and women (Sternang et al., 2014). Moreover, as suggested in a phenotypic study of Swedish twins, grip strength trajectories might be affected by different environmental factors in men and women (e.g. marital status had significant impact for men only, depression and dementia for women only) (Sternang et al., 2014).

Previous studies have estimated that the heritability of grip strength is approximately 50-60% (Frederiksen et al., 2002; Silventoinen et al., 2008), and a Swedish longitudinal study of grip strength found that the heritability was higher in men (75%) than in women (47%) (Finkel et al., 2003). Moreover, two studies based on Danish and Swedish twin data, respectively, have reported relatively constant heritability across the age range 45-96 years (Finkel et al., 2003; Frederiksen et al., 2002). Revisiting the Danish twin data using
age as a continuous variable, McGue et al established a slightly curvilinear heritability of grip strength across ages 45-96 years, the maximum heritability being observed in the youngest (approximately 60%) and reaching a minimum heritability of approximately 50% at age 70 years (McGue & Christensen, 2013). Only one longitudinal study has investigated the heritability of decline over a 9-year period (three follow-up assessments) of grip strength in Swedish twins aged 50-96 years at baseline and found that for neither men nor women did genes have a significant contribution to the age-related decline of grip strength (Finkel et al., 2003).

Despite the relatively substantial heritability of grip strength, few associations with particular genotypes have been reported. However, the literature does suggest that the APOE gene is associated with physical performance in aging populations. Thus, in a longitudinal study over 12 years, APOE ε2 carriers had less decline in grip strength than APOE ε3 carriers, whereas the decline of APOE ε4 carriers did not differ significantly from that of APOE ε3 carriers, however in right hand measurements only (Batterham et al., 2013). Two studies have reported a statistically non-significant tendency towards APOE ε2ε2/ε2ε3 being associated with lower grip strength and APOE ε3ε4/ε4ε4 being associated with greater grip strength compared to APOE ε3ε3/ε4ε2 (Vasunilashorn et al., 2013; Alfred et al., 2014). Another study has reported associations between the APOE gene and Activities of Daily Living (ADL) – a phenotype often used in aging studies and which partly captures muscle strength. The study demonstrated that in men APOE ε3ε3 decreased the risk of ADL disability and APOE ε2ε3 increased the risk of disability of Instrumental Activities of Daily Living (IADL); however, in women APOE ε4ε4 carriers had a significantly decreased risk of ADL disability (Kulminski et al., 2008) compared to APOE ε4ε4 non-carriers. This latter study demonstrates that APOE haplotypes might have different impact on physical decline in men and women.

The primary aim of this study was to establish whether grip strength trajectories were affected by gene-environment (GxE) interaction and, secondly, if the first test was confirmative, to examine whether the APOE gene could be a possible candidate gene for the GxE interaction. Since monozygotic (MZ) twins have all genes in common, within-pair differences cannot be ascribed to genetic effects or shared environmental factors, leaving non-shared environmental factors only. First, we tested whether differences in grip strength trajectories, obtained from growth curve modeling of maximum grip strength performance, exhibited evidence of a mixture distribution. Secondly, we tested whether the variability of MZ within-pair differences of grip strength trajectories differs as a function of APOE haplotype categories. Confirmative results of this test will provide evidence of GxE interaction, i.e., evidence that genes in general, or APOE haplotypes specifically, enhance or reduce the involvement of unspecified environmental factors on grip strength trajectories.

Methods

Participants
The sample comprised twin data from seven individual studies representing four countries: two from the United States, two from Sweden, one from Finland, and two from Denmark. Five studies had longitudinal grip strength measurements (Table 1). All seven studies are part of the Interplay of Genes and Environment across Multiple Studies (IGEMS) consortium (Pedersen et al., 2013).

**United States studies**

The two studies from the United States were the Vietnam Era Twin Study of Aging (VETSA) (Kremen et al., 2013) and the twin sample from Midlife Development in the United States (MIDUS) (Kendler et al., 2000). Both were longitudinal, but grip strength data were available from one occasion only. The VETSA study comprised male twin pairs aged 51-60 years at first assessment, and the age range of the twins from MIDUS, which included both sexes, was 34-82 years.

**Swedish studies**

Ascertainment of the two Swedish studies was based on records from The Swedish Twin Registry (Lichtenstein et al., 2002) and included the longitudinal studies Swedish Adoption/Twin Study of Aging (SATSA) (Pedersen et al., 1991) and the twins from the study Origins of Variance in the Oldest-Old (OCTO) (McClearn et al., 1997). Participants of the SATSA in-person tests were 39-88 years of age at first assessment and were reassessed at 3-year intervals and maximum seven times. The OCTO participants were 79-99 years of age at first assessment and were revisited a maximum of four times at 2-year intervals.

**Finnish study**

The participants of the Finnish Twin Study on Aging (FITSA) were recruited from the Finnish Twin Cohort (Tiainen et al., 2004). Selected on the basis of age and zygosity only, 414 same-sex female twin pairs from the Finnish Twin Study on Aging (FITSA) were recruited for clinical examination at age 63-76 years. Only pairs where both twins agreed to participate were invited for an examination. Survivors were invited for a second clinical examination three years later.

**Danish studies**

The Danish studies included the Longitudinal Study of Aging Danish Twins (LSADT) (Christensen et al., 1999) and the study of Middle-Aged Danish Twins (MADT) (Skytthe et al., 2013). Participants in these two studies were recruited from the Danish Twin Register which contains all identifiable twins born since 1870 (Skytthe et al., 2002). LSADT participants were 70-100 years and MADT participants were 45-68 years at first assessment. The LSADT study was initiated in 1995 and surviving participants, along with twins from younger birth cohorts, were invited for consecutive interviews every second year. Initially, the LSADT participants were same-sex twins aged 75+, but the inclusion age was progressively dropped to age 70 in 1999. Grip strength was not part of the battery until the 1999 survey. The MADT study comprised same-sex and opposite-sex twins who were visited in 1998, and surviving twins were invited to participate in a follow-up study 10 years later.
The total sample comprised 10,681 individual twins 34-99 years of age, including 1,724 same-sex MZ twin pairs with grip strength measurements; 1,141 of these pairs were genotyped for APOE (Table 1). All analyses were carried out separately for each sex as previous studies have demonstrated that heritability (Finkel et al., 2003) of grip strength and type of environmental factors influencing grip strength trajectories (Sternang et al., 2014) vary between sexes.

Measures

Grip strength

Grip strength was measured at in-person testing by trained interviewers; however, the protocols and the brand of the measuring devises differed among studies:

United States studies (MIDUS and VETSA)

In VETSA, grip strength was assessed using a JAMAR handheld dynamometer. The participants were seated in a study chair parallel to a table, resting one arm on the table while sitting with their back straight. The arm was positioned with the elbow flexed to 90 degrees and the wrist resting just off the end of the table. Participants were coached to push as hard as possible to obtain peak performance. The largest integer which the needle passed was recorded in kg. This was repeated, using alternating hands, starting with the dominant hand, until three trials were obtained for each hand.

In MIDUS, grip strength was measured (as part of MIDUS II) in six attempts (three on each hand) by a handheld dynamometer and always right hand first. The participant was instructed to support the elbow on a table, arm of chair or knee and squeeze as hard as possible until the measurement needle stopped moving.

Swedish studies (SATSA and OCTO)

In SATSA, grip strength was measured using a Collins dynamometer at sessions at a location convenient for the twin (Pedersen et al., 1991). The participants were placed in a seated position using a table as support for the elbow (Sternang et al., 2014) and had three trials on each hand.

In OCTO-twin, a Martin balloon dynamometer was used to measure grip strength at home-based interviews performed by nurses. The bulb of the dynamometer was adjusted to the hand size, and the arm rested on a table at a 45 degree angle (Proctor et al., 2006). The participants had three trials on each hand.

Finnish study (FITSA)

In FITSA, grip strength was measured using a dynamometer fixed to a chair. Maximal grip strength was measured at three to five attempts. The tests were done by trained physiotherapists (Tiainen et al., 2004).

Danish studies (MADT and LSADT)
In the two Danish studies, a handheld Smedley dynamometer was used and grip strength was measured three times on each hand during home-based interviews performed by trained lay-interviewers. The handle was adjusted to fit the size of the hand, and the participants were instructed to squeeze as hard as possible while holding their arm tight to the body and arm flexed in a 90 degree angle. The participant could choose a sitting or standing position during the test (Frederiksen et al., 2006).

In the seven studies, maximum grip strength measurements were obtained for each participant. Due to the differing procedures for grip strength measurements among the studies, all analyses were performed on standardized maximum grip strength measurements. The standardization was based on sex- and study-specific means and standard deviations from the first available waves in the respective studies (mean zero and standard deviation of 10).

**APOE-genotyping**

Genotyping of the *APOE* gene was performed in all studies except in MIDUS.

**United States studies (VETSA)**

In VETSA PCR and the HhaI, restriction digest methods were used to determine *APOE* genotypes (Schultz et al., 2008).

**Swedish studies (SATSA and OCTO)**

In the two Swedish studies, the two *APOE* markers (rs429358 and rs7412) were genotyped separately using Illumina GoldenGate assays (Reynolds et al., 2013).

**Finnish study (FITSA)**

The *APOE* genotypes were derived from SNP data obtained from genotyping on the Illumina HumanCoreExome chip, and subsequent imputation to 1000G.

**Danish studies (MADT and LSADT)**

In the Danish studies, genotyping was not performed on the total twin samples but only on randomly selected samples of the twin pairs. Genotyping of the *APOE* variants rs429358 and rs7412 were carried out using either custom-made primers and probes (LSADT), or predesigned TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, CA, USA) (MADT).

*APOE* haplotypes were grouped into three categories: *APOE* ε2ε2 and ε2ε3 (*APOE* ε2+), *APOE* ε3ε3, and *APOE* ε3ε4 and ε4ε4 (*APOE* ε4+), i.e., *APOE* ε2ε4 carriers were omitted from further analyses.

**Zygosity**
For most of the twins, the zygosity determination was based on twin responses to questions regarding similarity in physical appearance, a method whose validity has previously been shown to have an overall misclassification rate of less than 5% (Christiansen et al., 2003; Krueger & Johnson, 2002) For FITSA, VETSA, OCTO, and SATSA zygosity was confirmed by DNA analyses.

Analytic approach

Growth curve estimation

Features of longitudinal trajectories of grip strength were estimated using multilevel mixed linear regression models with full-information maximum likelihood estimation. The growth curve estimation was based on the total twin sample and Best Linear Unbiased Prediction estimates (BLUP’s) for intercept, and slope of the standardized grip strength measurements was estimated using age and age-squared (centered at 70 years) entered as fixed effects. The characteristics of the data, i.e., few measurement points on many individuals, did not allow for modelling of the random effect of age-squared, and therefore only linear effect of age was modelled in the random effects. This approach requires a minimum of grip strength measured at one occasion but models grip strength measurements at up to seven occasions. Hence, the intercept reflected the grip performance at age 70 and the linear slope, the ‘tilt’ of the curve, i.e., instantaneous linear rate of change at age 70. The slope parameter was set as missing for individuals who had grip strength measurement at one occasion only. Subsequent analyses were weighted (using the reciprocal standard error) BLUP estimates, resulting in greater weighting of cases with more longitudinal data than those with fewer points. The analyses were conducted on the untransformed weighted estimates as well as on the rank normalized weighted BLUP estimates to avoid spurious GxE interactions (Reynolds et al., 2007). We used Bloms’ rank-normalization method (Ludwig, 1961) i.e.

normalized estimates = inverse-normal[(ranked-estimates – 3/8)/(n-1/4)]

where n is the number of MZ pairs.

Analyses to evaluate evidence of GxE on grip strength trajectories were performed using MZ intra-pair methods that evaluate the possibility of mixture distributions of pair differences (Fisher, 1925) and test for variance homogeneity by genotype (Martin et al., 1983), as applied to longitudinal trajectory phenotypes (Reynolds et al., 2007). While growth curve modeling was based on all twins, the subsequent heterogeneity tests of within-pair differences and variance by APOE haplotypes were constrained to MZ twin pairs. Further description of these methods is provided below.

Heterogeneity test (Fisher)

In 1925 Fisher proposed a test for mixture of distribution based on differences within MZ twin pairs only (Fisher, 1925). Fisher’s test assumes a Gaussian distribution of the analyzed variable which induces the within-pair difference to follow a Gaussian distribution as well. A significant result of Fisher’s test suggests deviations from Gaussian distribution (i.e., the presence of more than a single distribution) of within-twin
pair differences. Since MZ twins share all genes, the variation of the within-pair differences can be attributed
to unshared environmental factors only. Hence, a significant result of Fisher’s test suggests that there are
multiple groups of MZ twins who show different responses to unspecified environmental factors. These
groups may be characterized by different genotypes, i.e., there is a GxE interaction. The formula for the test
statistic is
\[ t = \frac{(\bar{d}^2 - \frac{n}{2} \bar{d}^2)}{s}, \]
where \( \bar{d}^2 \) is the mean of the squared within-pair difference, \( \bar{d} \) is the mean of within-pair difference, \( s = \frac{\bar{d}^2}{\sqrt{n}} \) is the standard error, and \( n \) is the number of MZ twin pairs. The test statistic takes a \( t \)-distribution
with \( n - 1 \) degrees of freedom. Since \( t \) is expected to be positive, we used a one-sided \( t \)-test.

**Variance homogeneity test**

Among MZ twin pairs only, we performed Bartlett’s test to compare the variability of within-twin pair
differences of weighted BLUP estimates in three \( APOE \) haplotype categories \( APOE \varepsilon2^+ \) (i.e. \( APOE \varepsilon2\varepsilon2 \) or
\( \varepsilon2\varepsilon3 \)), \( APOE\varepsilon3\varepsilon3 \), and \( APOE \varepsilon4^+ \) (i.e. \( APOE \varepsilon3\varepsilon4 \) or \( \varepsilon4\varepsilon4 \)). This test was performed on untransformed as
well as rank-normalized weighted BLUP estimates. Moreover, the test was performed on Winsorized estimates (i.e., outliers more than three SDs away from the mean were replaced with values equivalent to
three SDs from the mean) to reduce the risk of significant results caused by outliers. Significant
heterogeneity indicates that particular \( APOE \) haplotypes may be more or less sensitive to environmental
factors, i.e., that environmental factors interact with haplotypes of the \( APOE \) gene.

Stata version 13 (College Station TX, 2013) was used for all statistical analyses.

**Results**

Summary statistics for the 7 studies are presented in Table 1. Locally weighted regression curves, separately
by study, for standardized grip strength on age are shown in Figures 1 (men) and 2 (women).

Haplotype distribution by nationality and test for Hardy-Weinberg equilibrium of \( APOE \) haplotypes (online
calculator: http://www.had2know.com/academics/hardy-weinberg-equilibrium-calculator-3-alleles.html) are
given in Table 2. While no deviance from Hardy-Weinberg equilibrium was observed in the Swedish,
Finnish, and Danish data, there was evidence of Hardy-Weinberg disequilibrium of the \( APOE \) genotypes in
the United States data (\( p=0.01 \)). However, individual test of the two SNPs did not result in any deviation
from the Hardy-Weinberg equilibrium (rs429358: \( p=0.79 \), rs7412: \( p=0.05 \)).

**Growth curves**
Age- and study adjusted mean of growth curve parameters for all twins by APOE haplotypes and sex are reported in Table 3. Though not statistically significant, in men the mean of the intercept was slightly lower in the APOE ε2+ haplotypes, whereas in women the direction was opposite, i.e. APOE ε2+ carriers had the highest intercept, and APOE ε4+ carriers had the lowest intercept. There were no differences in the mean of slopes in the three haplotype categories in men or women.

**Heterogeneity tests**

Fisher’s test for mixture of distribution adjusted for age and study was highly significant in both sexes for the trajectory features before as well as after rank-normalization. The results indicated that within-pair (MZ pairs only) differences of grip strength trajectories deviated significantly from a single Gaussian distribution; thus the analyses indicated that there are different groups whose grip strength trajectories showed different responses to unspecified environmental factors (Table 4). Hence, we tested whether APOE haplotypes might index the groups that vary in environmental sensitivity.

Table 5 reports Bartlett’s test for equal variances of within MZ twin pair age and study adjusted differences of grip strength trajectory features in three APOE categories stratified by sex. Significant heterogeneity indicates that particular APOE haplotypes may be more or less sensitive to environmental factors, i.e., that environmental factors interact with haplotypes of the APOE gene. The results showed a trend of increasing variability of the trajectory features across APOE haplotype categories (from ε2+ to ε4+) in women. However, in men, the largest variability of the intercept was observed in APOE ε3ε3, and the variability of the slope was similar in the APOE ε3ε3 and APOE ε4+ groups. The only statistically significant result was found for the slope (p<0.01) in women, but this was not retained in the analyses of the Winsorized or rank-normalized estimates, which indicates that the significance was driven by outliers. However, statistical strength was retained when the intercept estimate was rank normalized (p=0.04). Notably, while Winsorization impacted the significance of the tests, it had little impact on the variances (results not shown).

**Discussion**

In the present study we examined grip strength trajectories in a large sample of twins pooled from seven surveys across four countries. The differences in mean levels of the trajectory features by APOE haplotype categories were small and statistically not significant in general. We found evidence of GxE interaction on the trajectory features. Moreover, our results suggest that the APOE gene might be a candidate gene for the GxE interaction. To our knowledge, this is the first study to address the question of GxE interaction in grip strength trajectories.
Previous studies of grip strength and APOE haplotypes found statistically non-significant tendencies towards lower grip strength in APOE ε2+ over APOE ε3ε3 to APOE ε4+ (Vasunilashorn et al., 2013; Alfred et al., 2014). In our study, we found that among men the intercept at age 70 was slightly lower in the APOE ε2+ group, but highest in the APOE ε3ε3 group. In women, the APOE ε2+ carriers showed greater and the APOE ε4+ carriers lower grip strength levels than APOE ε3ε3 carriers. Thus, our results do not confirm the tendencies found in previous studies. This might be due to the fact that we stratified the analyses by sex, whereas previous studies adjusted for sex, thereby possibly masking different directional trends in men and women. We did not observe any differences in linear slope across APOE haplotype categories. Consequently, despite the large sample sizes of our study, we did not strong evidence of any association between the APOE gene and mean of trajectory features.

The results from Fisher’s test for mixture distribution of within-MZ-twin pair differences of the trajectory features (i.e. linear slope and intercept at age 70) demonstrated general evidence of GxE interaction for men and women. It is possible that the missing heritability of change in grip strength (Finkel et al., 2003) was obscured by the existence of GxE interaction since this would contribute to the unique environment and not the genetic variance components in heritability analyses.

Our study offered consistent, though not statistically significant, evidence that the variances of within-pair differences in trajectory parameters in MZ twins were smaller in the APOE ε2+ haplotype category than in the other categories. These results could suggest that, compared to other APOE haplotypes, carriers of the APOE ε2+ haplotypes may be less sensitive to (unspecified) unshared environmental factors, i.e., that there was an interaction between the APOE gene and unspecified environmental factors affecting the grip strength trajectories. This interpretation may be in line with previous studies suggesting that the APOE gene interacts with environmental factors on some phenotypes related to grip strength. Thus, in a study of earthquake victims it was shown that, one year after the earthquake, APOE ε4+ haplotypes had lower levels of self-rated health, mobility and IADL (Daly & MacLachlan, 2011), and in another study of male twins, lower total cerebral brain volume was associated with worse physical performance (composite of walking speed, balance, and chair stand) in APOE ε4+ carriers than in APOE ε4 non-carriers (Carmelli et al., 2000).

Grip strength is a phenotype that has been associated primarily with late-life, age-related health outcomes. However, several studies have demonstrated that grip strength declines throughout midlife to late-life. Therefore, the growth curve modelling in our study was based on grip strength measurements of twins in a wide age range (34-99 years) applying curvilinear main effects of age. Thus, we took advantage of the wide age range to model the decline of grip strength throughout mid- to late-life. However, this approach also relied on the assumption that the GxE interaction was conserved across the age-range. We repeated the analyses stratified in two groups (those who were less than age 70 at intake and those who were age 70 or more at intake) which lowered the statistical power but the trends across APOE groups were preserved (results not shown).
The large sample of informative MZ twins is a major strength of the present study. Our analytical approach is powerful since it controls for genetic influences and any common environmental influences. The differing protocols for grip strength measurements in the various studies were a limitation of our study; therefore grip strength was standardized separately by study prior to growth curve modelling. Secondly, apart from three studies (SATSA, OCTO, and LSADT), the number of possible measurement occasions was less than three which did not allow us to estimate the individual differences in the quadratic growth curve parameter for acceleration or deceleration of decline. Hardy-Weinberg Equilibrium of the APOE gene was not confirmed in VETSA. However, performing a chi-squared test on the two single SNP’s did not provide any evidence of a violation of the Hardy-Weinberg Equilibrium (both p > 0.05). Moreover, Hardy-Weinberg Equilibrium was not violated in the total sample. Thus, we did not expect the deviance from Hardy-Weinberg Equilibrium in the VETSA study to introduce any bias. Last, tests of equality of variances are beset by low power (e.g., (Martin et al., 1983)), which would have been of particular concern for tests of the linear slopes.

The analyses in our study were based on the maximum of the attempted grip strength measures. Alternatively, as is most commonly described in the literature, the average of the attempts could have been used. Previous studies of the validity of grip strength have demonstrated that grip strength decreased by each attempt suggesting increasing fatigue (Abizanda et al., 2012; Watanabe et al., 2005). However, allowing the participant to rest 1 minute between each attempt gave stable outcomes of the attempts (Watanabe et al., 2005). Since each survey in our study has its own protocol for measuring grip strength, but none of the protocols specify any recommended rest interval between the attempts, we expected the maximum grip strength to be more reliable across studies than the average grip strength. Moreover, although it is possible to underestimate maximum grip strength, if maximum effort is not used, it is exceedingly difficult to conceive of a way that an individual could produce a grip strength result that was higher than his or her true maximum.

Further analyses on larger sample sizes should be performed to examine the possibility of an interaction between APOE (as well as other genes) and unspecified environmental factors on grip strength trajectories. Search for specific environmental factors whose effect on grip strength trajectories are modified by the APOE gene (or other genes) could be selected among those environmental factors that have been found to affect grip strength (e.g. smoking, socioeconomic status, education, early malnutrition, stature, strenuous work, and diseases) and, as suggested by Sternäng et al (Sternang et al., 2014), different environmental factors may be involved for men and women.

Acknowledgements

IGEMS is supported by the National Institutes of Health grant No. R01 AG037985. SATSA was supported by grants R01 AG04563, R01 AG10175, the MacArthur Foundation Research Network on Successful Aging, the Swedish Council For Working Life and Social Research (FAS) (97:0147:1B, 2009-0795) and
Swedish Research Council (825-2007-7460, 825-2009-6141). OCTO-Twin was supported by grant R01 AG08861. The Danish Twin Registry is supported by grants from The National Program for Research Infrastructure 2007 from the Danish Agency for Science Technology and Innovation, the Velux Foundation and the US National Institute of Health (P01 AG08761). VETSA was supported by National Institute of Health grants NIA R01 AG018384, R01 AG018386, R01 AG022381, and R01 AG022982, and, in part, 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 Registry. The MIDUS study was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development and by National Institute on Aging Grant AG20166. FITSA was supported by grants from the Academy of Finland (69818) and the Finnish Ministry of Education and Culture (120/722/2003).


College Station TX: StataCorp LP. (2013). Stata: Release 13. Statistical software. TX: StataCorp LP: College Station.


<table>
<thead>
<tr>
<th>Study</th>
<th>Number of individual twins in each wave</th>
<th>Male (%)</th>
<th>Age range (median) at baseline</th>
<th>Max (median) number of grip strength measurements</th>
<th>Number of MZ pairs</th>
<th>Number of MZ pairs with APOE genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>VETSA (USA)</td>
<td>1,215</td>
<td>100%</td>
<td>51-60 (54)</td>
<td>1 (1)</td>
<td>311</td>
<td>308</td>
</tr>
<tr>
<td>MIDUS (USA)</td>
<td>379</td>
<td>41%</td>
<td>34-82 (53)</td>
<td>1 (1)</td>
<td>81</td>
<td>----</td>
</tr>
<tr>
<td>SATSA (Sweden)</td>
<td>851; 741; 646; 468; 322; 232; 141</td>
<td>41%</td>
<td>39-88 (63)</td>
<td>7 (4)</td>
<td>153</td>
<td>133</td>
</tr>
<tr>
<td>OCTO (Sweden)</td>
<td>640; 511; 383; 274; 190</td>
<td>34%</td>
<td>79-99 (82)</td>
<td>5 (3)</td>
<td>127</td>
<td>113</td>
</tr>
<tr>
<td>FITSA (Finland)</td>
<td>434; 308</td>
<td>0%</td>
<td>63-79 (69)</td>
<td>2 (1)</td>
<td>103</td>
<td>101</td>
</tr>
<tr>
<td>MADT (Denmark)</td>
<td>4,276; 2,358</td>
<td>51%</td>
<td>45-77 (56)</td>
<td>2 (2)</td>
<td>657</td>
<td>386</td>
</tr>
<tr>
<td>LSADT (Denmark)</td>
<td>2,886; 2,121; 1,585; 882</td>
<td>45%</td>
<td>70-97 (75)</td>
<td>4 (3)</td>
<td>292</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>10,681</td>
<td>51%</td>
<td>34-99 (66)</td>
<td>7 (1)</td>
<td>1,724</td>
<td>1,141</td>
</tr>
</tbody>
</table>

*including broken pairs

VETSA Vietnam Era Twin Study of Aging; MIDUS Midlife Development in the United States; SATSA Swedish Adoption/Twin Study of Aging; OCTO Origins of Variance in the Oldest-Old; FITSA Finnish Twin Study of Aging; MADT Middle Aged Danish Twins; LSADT Longitudinal Study of Aging Danish Twins.
Table 2 Frequency count and test for Hardy-Weinberg equilibrium of the total samples of APOE haplotypes by country.

<table>
<thead>
<tr>
<th>Study</th>
<th>ε2ε2 (p-value)</th>
<th>ε2ε3</th>
<th>ε3ε3 (p-value)</th>
<th>ε2ε4</th>
<th>ε3ε4</th>
<th>ε4ε4 (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA a</td>
<td>2 (0.2%)</td>
<td>110</td>
<td>508 (58.5%)</td>
<td>36</td>
<td>193</td>
<td>19 (2.2%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden b</td>
<td>8 (0.8%)</td>
<td>139</td>
<td>559 (55.8%)</td>
<td>32</td>
<td>240</td>
<td>23 (2.3%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland c</td>
<td>1 (0.3%)</td>
<td>24</td>
<td>210 (64.0%)</td>
<td>6</td>
<td>82</td>
<td>5 (1.5%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark d</td>
<td>14 (0.6%)</td>
<td>296</td>
<td>1,269 (55.7%)</td>
<td>68</td>
<td>568</td>
<td>65 (2.9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25 (0.6%)</td>
<td>569</td>
<td>2,546 (56.9%)</td>
<td>142</td>
<td>1,083</td>
<td>112 (2.5%)</td>
</tr>
</tbody>
</table>

Note: One individual from each MZ pair is left out. An online calculator for testing Hardy-Weinberg equilibrium of three alleles was used (http://www.had2know.com/academics/hardy-weinberg-equilibrium-calculator-3-alleles.html)

a VETSA (Vietnam Era Twin Study of Aging)
b SATSA (Swedish Adoption/Twin Study of Aging) and OCTO (Origins of Variance in the Oldest-Old)
c FITSA (Finnish Twin Study of Aging)
d MADT (Middle Aged Danish Twins) and LSADT (Longitudinal Study of Aging Danish Twins)
Table 3 Mean of trajectory features, adjusted for study and age at first assessment, for all twins broken by APOE haplotype and sex.

<table>
<thead>
<tr>
<th>APOE ε2+</th>
<th>APOE ε3ε3</th>
<th>APOE ε4+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean (95%CI)</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept estimates</td>
<td>378</td>
<td>-0.19 (-0.39;0.02)</td>
</tr>
<tr>
<td>Linear slope estimates</td>
<td>171</td>
<td>0.03 (-0.06;0.12)</td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept estimates</td>
<td>371</td>
<td>0.17 (-0.02;0.35)</td>
</tr>
<tr>
<td>Linear slope estimates</td>
<td>281</td>
<td>0.01 (-0.03;0.06)</td>
</tr>
</tbody>
</table>

Note Estimates are weighted Best Linear Unbiased Prediction estimates of trajectory features (see text)

Note The APOE ε2+ consists of APOE ε2ε2 and APOE ε2ε3 and the APOE ε4+ consists of APOE ε3ε4 and APOE ε4ε4. APOE ε2ε4 is excluded from the analyses

*p-values (ANOVA analyses) for equality of means across APOE groups*
Table 4 Fisher’s heterogeneity test for mixture distribution of within MZ twin pair differences of trajectory features broken by sex. The analyses are adjusted for age at first assessment and study.

<table>
<thead>
<tr>
<th></th>
<th>Intercept Estimates</th>
<th>Linear slope Estimates</th>
<th>Rank-normalized estimates</th>
<th>Rank-normalized estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Z&lt;sup&gt;b&lt;/sup&gt;</td>
<td>P&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td>917</td>
<td>8.72</td>
<td>&lt;0.001</td>
<td>4.51</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td>807</td>
<td>5.00</td>
<td>&lt;0.001</td>
<td>4.23</td>
</tr>
</tbody>
</table>

*Note* Estimates are weighted Best Linear Unbiased Prediction estimates of trajectory features (see text)

<sup>a</sup> Number of MZ pairs

<sup>b</sup> Fisher’s test statistic for deviance from normal distribution \( t = \left( \frac{\bar{d}^2 - \frac{\pi}{2} \bar{d}^2}{s} \right) / s \), where \( \bar{d}^2 \) is the mean of the squared within-pair difference, \( \bar{d} \) is the mean of within-pair difference, \( s = \frac{\bar{d}^2}{n} \) 0.532 is the standard error, and \( n \) is the number of MZ twin pairs.

<sup>c</sup> One-sided t-test; n-1 degrees of freedom
Table 5 Bartlett’s test for equal variances of within-pair differences, MZ twins only, of trajectory features in categories of APOE haplotypes. The analyses are adjusted for age at first assessment and study.

<table>
<thead>
<tr>
<th></th>
<th>APOE category</th>
<th>N</th>
<th>Variance of trajectory estimates (95%CI)</th>
<th>Trajectory estimates</th>
<th>Winsorized estimates&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Rank-normalized estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>ε2ε2 or ε2ε3</td>
<td>86</td>
<td>1.85 (1.26;2.45)</td>
<td>0.32</td>
<td>0.58</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>ε3ε3</td>
<td>347</td>
<td>2.41 (1.89;2.92)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ε3ε4 or ε4ε4</td>
<td>183</td>
<td>2.23 (1.68;2.77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear slope</td>
<td>ε2ε2 or ε2ε3</td>
<td>32</td>
<td>0.19 (0.11;0.28)</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>ε3ε3</td>
<td>115</td>
<td>0.37 (0.25;0.49)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ε3ε4 or ε4ε4</td>
<td>62</td>
<td>0.37 (0.20;0.54)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>ε2ε2 or ε2ε3</td>
<td>69</td>
<td>1.75 (1.09;2.42)</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>ε3ε3</td>
<td>281</td>
<td>2.63 (2.07;3.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ε3ε4 or ε4ε4</td>
<td>139</td>
<td>2.94 (2.12;3.75)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear slope</td>
<td>ε2ε2 or ε2ε3</td>
<td>43</td>
<td>0.16 (0.03;0.30)</td>
<td>&lt;0.01</td>
<td>0.49</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>ε3ε3</td>
<td>172</td>
<td>0.21 (0.15;0.28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ε3ε4 or ε4ε4</td>
<td>82</td>
<td>0.36 (0.06;0.65)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>Note</sup> Estimates are weighted Best Linear Unbiased Prediction estimates of trajectory features (see text)

<sup>a</sup> Winsorization: absolute values greater than 3 SD are set to +/- 3SD respectively.
Captions (figures)

Figure 1. Locally weighted regression curves of the standardized grip strength on age for all men in the total sample as well as in the single studies. The thick gray curves are for the total sample.

Figure 2. Locally weighted regression curves of the standardized grip strength on age for all women in the total sample as well as in the single studies. The thick gray curves are for the total sample.
Figure 1
Figure 2