Behavior Genetics, Vol. 33, No. 2, March 2003 (© 2003)

Age Trajectories of Genetic Variance in Physical Functioning: A Longitudinal Study of Danish Twins Aged 70 Years and Older

Kaare Christensen,[1,4] Henrik Frederiksen,[1] James W. Vaupel,[2] and Matt McGue[3]

Genetic-evolutionary theories of aging predict that the genetic variance for fitness traits increases with age, while epidemiological-gerontological theories predict an increase in the environmental variance for most traits. In this study we examine the age trajectories of the genetic and environmental variance in physical functioning in a sample of 4731 Danish twins aged 70+ who are being followed longitudinally every second year with up to four assessments completed. A biometric growth model (Neale and McArdle, 2000) was applied to a validated physical ability score. The model included an overall level effect, a rate of linear change effect, and residual effects. The best-fitting model was a sex-specific model including additive genetic and nonshared environmental factors affecting level and rate of change and only nonshared environmental factors affecting the wave-specific levels. For both sexes there is an approximate doubling of both the total variance and the genetic variance in the physical ability score over the four waves and, hence, a rather stable heritability. However, the heritability is approximately .10 for males and .30 for females in all four waves. The heritability of level and slope showed a similar pattern: .11–14 in males and .35–.39 in females. The increase in both additive genetic variance and environmental variance is in agreement with genetic-evolutionary and epidemiological-gerontological theories of aging, respectively. The present study suggests that overall level of strength may be a better phenotype for future molecular genetic studies on physical functioning in the elderly than rate of change, because rate of change is vulnerable to sample attrition due to mortality and dropout and because four waves were needed to be able to detect a heritability for rate of change of the same magnitude as the heritability for level of physical functioning.

KEY WORDS: Aging; twins; genetics; evolution; physical abilities; rate of change.

INTRODUCTION

The maintenance of adequate physical functioning is, together with intact cognitive abilities, key to maintaining an independent and active late life. With age, physical functioning becomes increasingly variable: A substantial proportion of the elderly suffer from severe loss of abilities, while others maintain functioning similar to that of many middle-aged individuals or experience only small or moderate declines. Cross-sectional twin studies suggest that the basis for variation in both late-life physical and cognitive functioning is attributable to both environmental and genetic factors and that genetic factors may become increasingly important with age (Christensen *et al.*, 2000; McClearn *et al.*, 1997; McGue and Christensen, 2001).

Major theories of aging, such as genetic-evolutionary theories and epidemiological-gerontological theories, make predictions on the age trajectory of genetic

[1] The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, DK-5000 Odense, Denmark.
[2] Max Planck Institute for Demographic Research, Rostock, D-18057, Germany.
[3] Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA.
[4] To whom correspondence should be addressed at Epidemiology Unit, Institute of Public Health, University of Southern Denmark, Sdr. Boulevard 23A, DK-5000 Odense C, Denmark. e-mail: kchristensen@health.sdu.dk

0001-8244/03/0300-0125/0 © 2003 Plenum Publishing Corporation

and environmental contributors to variance in physical functioning. The genetic-evolutionary theories of aging are based on the proposition that weak natural selection late in life (postreproductive) leads to accumulation of mutations with deleterious late-life effects. These models therefore predict that additive genetic variance will increase with age for fitness traits (i.e., traits associated with reproductive success). This prediction is made by both variants of the genetic-evolutionary hypothesis: the mutation accumulation theory, according to which the mutations are neutral in early life, and the antagonistic pleiotropy theory, according to which the mutations are beneficial in early life but deleterious in late life (Charlewood and Hughes, 1996).

The epidemiological-gerontological theories of aging, on the other hand, consider "wear and tear" as the central mechanism in aging, especially oxidative reactions on a molecular level, but also other chemical reactions, radiations, and so forth (Campisi, 2001). This accumulation of damage over time will lead to an increase in the environmental variance with age. The epidemiological-gerontological theories are often not very explicit in predicting changes in genetic variance, but it is often assumed that the heritability will decrease with age due to increases in environmental variance (Harris *et al.*, 1992). It is important to note that the genetic-evolutionary theories and the epidemiological-gerontological theories are not mutually exclusive—both genetic and environmental variance can increase with age.

Tests of the theories and their predictions about the age trajectories of genetic variance have been performed using both animal and human data. There is a comprehensive experimental literature on *Drosophila melanogaster* addressing the genetic-evolutionary theories of aging. Although results have been conflicting, the overall pattern of findings supports both the mutation accumulation and antagonist pleiotropy theories (Patridge, 2001). In humans, there have been fewer empirical tests of genetic-evolutionary theories of aging. Such studies are difficult primarily because survival data in which the effects of common environment and genetic factors can be disentangled are not readily available and, secondly, because fertility is highly regulated in most human populations (Hughes and Burleson, 2000).

In this study we examine the age trajectory of the genetic variance in physical functioning in a sample of Danish twins aged 70+ who are being followed longitudinally every second year with up to four assessments completed. Our analysis of the three first waves of data (Christensen *et al.*, 2002) revealed that the level of physical abilities in females is moderately heritable (.3–4), while the corresponding heritability in males is nonsignificant. In addition, the genetic contribution to rate of change in physical abilities was not statistically significant. A maximum of three assessments covering a span of no more than 4 years, however, may not be long enough for changes in physical functioning to be of sufficient magnitude to allow reliable detection of genetic influences. The present study benefits from the increased power gained by a fourth assessment (Bryk and Raudenbush, 1987). It also allows us to test the predictions of the above-mentioned major theories of aging directly by assessing age changes in genetic and environmental components of variance, a test that was not undertaken in the original study.

METHODS

Sample

The sample is based on participants in the Longitudinal Study of Aging Danish Twins (LSADT) (Christensen *et al.*, 1999). LSADT began in 1995 with the assessment of members of like-sex twin pairs born in Denmark prior to 1920 (i.e., at least 75 years old at the beginning of 1995). The surviving members of the initial cohort were followed up every 2 years in 1997, 1999, and 2001. Additional cohorts were also added at the 1997, 1999, and 2001 assessments and subsequently followed at 2-year intervals. Twin pairs in which both members were alive and born between 1920 and 1923 (i.e., at least age 73 years at the beginning of 1997) were added in 1997; twins born 1920 to 1928 (i.e., at least 70 years old at the beginning of 1999) were added in 1999; and twins born between 1929 and 1930 (i.e., at least 70 years old at the beginning of 2001) were added in 2001. An overview of the LSADT cohort-sequential design is given in Fig. 1.

The sample ascertainment procedures are described briefly here; additional details are given in Christensen *et al.* (1999). The sample was drawn from the older cohorts of the Danish Twin Registry, which includes twins born between 1870 and 1910 and like-sex twin pairs born between 1911 and 1930 in Denmark (Hauge, 1981; Holm, 1983). The residence of all eligible and surviving twins was determined from the Danish Central Person Registry just prior to the scheduled time of assessment. Participation rates among survivors in a given wave ranged from approximately 70% to 80%. Christensen *et al.* (1999) report a comparison of participants and nonparticipants in the 1995 survey,

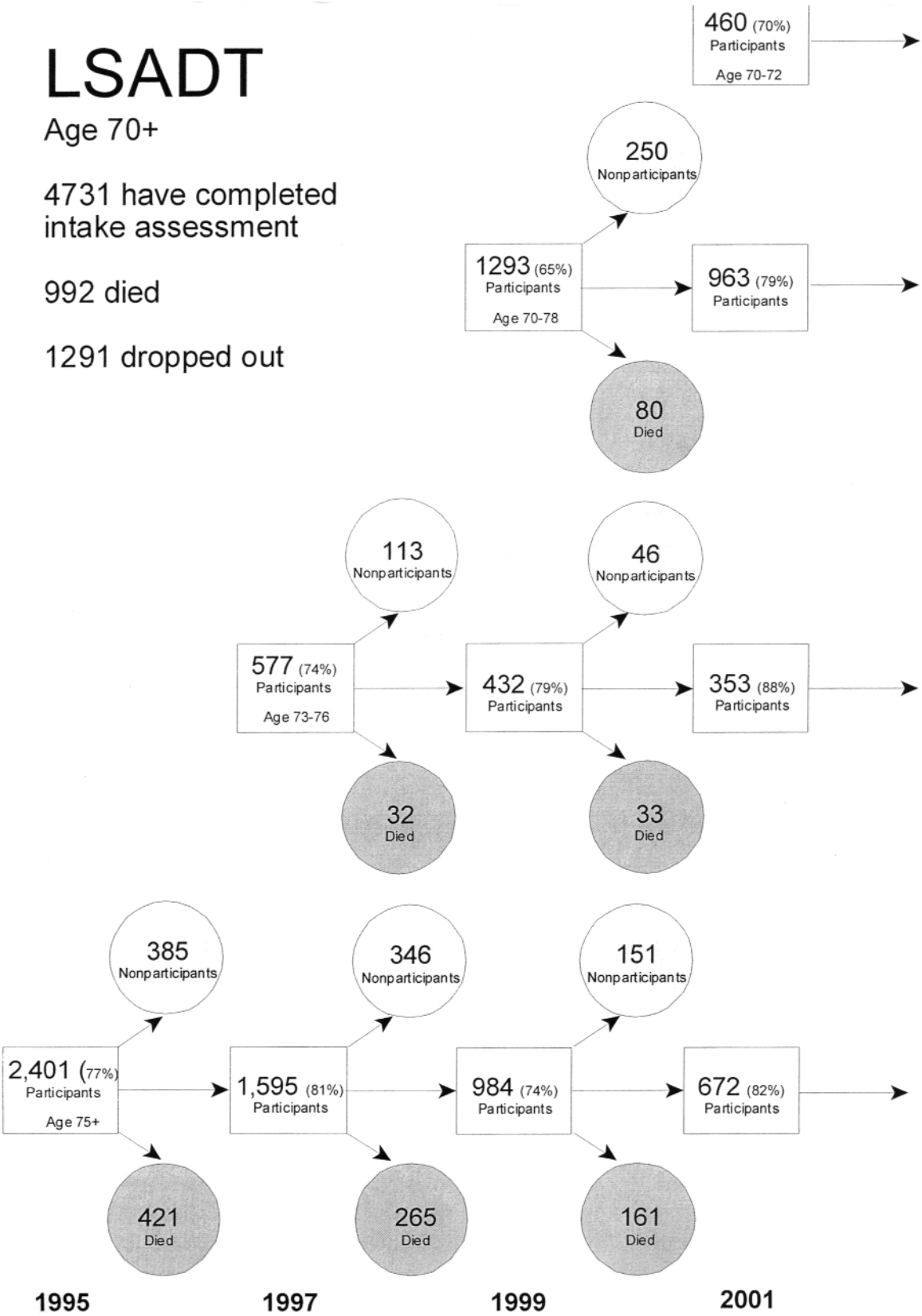

Fig. 1. Overview of the Longitudinal Study of Aging Danish Twins (LSADT), which uses a cohort-sequential design.

showing that participants were slightly but significantly more likely to be male and to have been hospitalized between 1977 and 1994 than were nonparticipants but did not differ significantly from nonparticipants in terms of age and twin zygosity.

A total of 4731 individual twins have completed an LSADT intake assessment, either through in-person interview or by proxy. All but 28 (0.6%) of the 4731 participating twins had a valid intake assessment of physical functioning. Because for the most part twins were selected without regard to their co-twin's mortality (the exception being the 1920 to 1923 birth cohorts, where both twins needed to be alive at the start of 1997 to qualify for participation), many of the participating twins did not have a surviving co-twin who participated in the survey. Among intake participants, there were 2304 individual twins from 1152 intact twin pairs (451 monozygotic [MZ], 661 dizygotic [DZ], 19 of unknown zygosity, and 21 opposite sex) and 2427 single twins.

The analysis reported here is based on the 1112 like-sex twin pairs with known zygosity. Table I gives the number of twins who completed the physical abilities assessment at each of the four waves. A total of 1105 twin pairs and 7 individual twins completed the assessment of physical functioning at intake. Because of the cohort sequential nature of the design and twin mortality, the number of twin pairs available for analysis drops off rapidly with each successive wave. Also given in the table is the mean age of the samples at each assessment wave. Although the interval between assessment waves was fixed at 2 years, the difference in the average age of the sample between successive assessments differs by more than 2 years because the minimum age for eligibility was reduced in each successive survey (i.e., the minimum age was 75 years in 1995 and 70 years in 2001).

Procedures

The assessments were usually completed in the twins' primary residence between February and April of the scheduled year. Assessments were administered by approximately 100 interviewers from the Danish National Institute of Social Research, which has substantial experience interviewing the elderly (Kjøller, 1995; Platz, 1989, 1990). Interviewers completed an extensive training program in the 2 months prior to the start of the survey and were closely monitored during the survey period. Different interviewers interviewed the two members of intact twin pairs, and many interviewers participated in multiple waves. Interviews lasted approximately 60 to 75 minutes and included assessment of physical health, functional status, background demographics, cognitive functioning, and depression symptomatology.

Measures

Functional Abilities

The instrument (Avlund) used in this study has previously been validated in Denmark (Christensen *et al.*, 2000; Avlund, Davidson, and Schultz-Larsen 1995; Schultz-Larsen, Avlund, and Kreiner, 1992). The assessment of functional abilities was based on self-report, which has generally been found to be reliable and valid (Kane and Kane, 1981; Jette, 1987). The Avlund instrument has been described in detail previously and has been shown to discriminate levels of functional abilities among community-dwelling elderly through questions about tiredness and the need for personal assistance in relation to functional abilities (Avlund, Davidson, and Schultz-Larsen, 1995; Schultz-Larsen, Avlund, and Kreiner, 1992). The Avlund instrument

Table I. Twins Completing the Physical Abilities Assessment at Each Wave of the Longitudinal Study of Aging Danish Twins

	MZ twins		DZ twins		Total		Mean age (years)
Assessment	Pairs	Singles	Pairs	Singles	Pairs	Singles	(SD)
Intake	449	2	656	5	1105	7	76.0 (4.6)
Follow-up 1	252	87	333	193	585	280	78.3 (4.1)
Follow-up 2	102	89	122	174	224	263	82.1 (3.8)
Follow-up 3	40	51	37	110	77	161	85.6 (3.7)

Note: The sample is derived from the 1112 like-sex twin pairs with known zygosity. Entries are the numbers of twin pairs and nonmatched single members of twin pairs who completed the physical abilities assessment at each assessment wave. Decline in sample size across successive waves owes primarily to the cohort-sequential design and mortality.

was extended to include assessment of need for equipment or aids in relation to functional abilities, based on results showing that equipment and aids can improve functional abilities among the elderly (Manton, Corder, and Stallard, 1993). All the items from the Katz index of Activities of Daily Living were included (Katz *et al.*, 1963), as well as questions about the ability to see, hear, and carry out more demanding activities such as running (Christensen *et al.*, 2000).

To identify meaningful quantitative subscales, the 26 items were factor analyzed in the total LSADT-95 twin sample. All items were rated on a 1 to 4 scale, with the response options as follows: 4 = can do without fatigue, 3 = can do with fatigue or minor difficulties, 2 = can do with aid or major difficulties, 1 = cannot do. In the factor analyses, three factors had an eigenvalue of more than one, but few of the items loaded on the third factor. Therefore, a two-factor solution was adopted. The first factor loaded highest on items dealing with ability to walk, run, climb stairs, and carry weights (including 11 categories "Walk around in the house"; "Walk up and down stairs one floor"; "Walk up stairs to the second floor"; "Able to get outdoors"; "Able to walk 400 meters without resting"; "Do light exercise"; "Do hard exercise"; "Walk in nice weather for ½ to 1 hour"; "Walk in bad weather for ½ to 1 hour"; "Run 100 meters"; "Carry 5 kilos") and was interpreted to reflect a dimension of strength. The second factor loaded highest on items dealing with ability to dress and wash oneself and get in and out of bed (including 11 categories "Get up from a chair and a bed"; "Able to go to the toilet"; "Wash upper part of body"; "Wash lower part of body"; "Wash hair"; "Dress upper part of body"; "Dress lower part of body"; "Take socks and shoes on and off"; "Comb hair"; "Cut toenails"; "Cut fingernails") and was interpreted to reflect a dimension of agility. Scores for the two dimensions were calculated by taking the average response of items loading highest on the factor or having been judged to be relevant for that dimension. The internal consistency reliability estimate for the strength scale was .93 in both the male and female samples for both the in-person and the proxy interviews. The reliability estimates for the agility scale were also the same for males and females and equaled .91 for the in-person interview and .93 for the proxy interview. These values indicate very reliable scales. The correlation was .77 between the strength and the agility scales. The later surveys yielded very similar results. In this study we focus on the strength scale, since agility scale scores varied little among younger participants and so the scale was dropped from the assessment in 1999.

Strength scores were corrected for the effects of gender and intake age prior to biometric analysis of the twin data by subtracting from each individual's observed strength score the appropriate age-gender group mean. Eight intake age-gender groups were formed for this purpose (two genders by four intake age-groups, 79 years and younger, 80–84 years, 85–89 years, and 90 years and older). Although age and gender effects can be modeled directly in the growth model analyses (McArdle *et al.*, 2002), we elected to account for these effects prior to the growth-curve analysis because of the small size of some of the gender-age combinations (e.g., nonagenarian males). Correlations among the age-gender adjusted intake and follow-up strength scores across the successive waves of assessment are given in Table II. As expected, the correlations decrease as a function of retest interval, with assessments separated by a single wave correlated .66–80, assessments separated by two waves correlated .57–.68, and the assessments separated by three waves correlated .55 (Table II).

Zygosity

Twin zygosity was based on a self-report questionnaire completed by members of the Danish Twin Registry when the sample was predominantly in middle age. The questionnaire method used to establish zygosity in the Danish Twin Registry has been validated against methods based on genetic markers and found to result in error rates of less than 5% (Hauge, 1981).

Statistical Procedures

Biometric analysis of twin data (Neale and Cardon, 1992) assumes that the phenotypic variance in a quantitative trait can be decomposed into components attrib-

Table II. Retest Correlations for Strength Scores for 2-Year Retest Intervals in Danish Twins Aged 70+

	Intake	Follow-up 1	Follow-up 2	Follow-up 3
Intake	—			
Follow-up 1	.76 $N = 2976$	—		
Follow-up 2	.68 $N = 1318$	.80 $N = 1318$	—	
Follow-up 3	.55 $N = 660$	.57 $N = 660$	.66 $N = 665$	—

Note: Correlations are based on age-gender adjusted scores and are all significantly greater than zero at $p < .001$.

utable to additive genetic effects (denoted as A), shared environmental effects (C, those environmental effects that are shared by reared-together relatives), and nonshared environmental effects (E, those environmental effects that differ for relatives). Under the additional assumption that shared environmental effects contribute equally to the similarity of MZ and DZ twins (i.e., the Equal Environmental Similarity Assumption), the separate contributions of these three variance components are estimable because MZ twins share 100% of the additive genetic effects, whereas DZ twins share on average only 50% of these effects. The Equal Environmental Similarity Assumption has received consistent support in the behavioral genetic literature (Plomin *et al.*, 2001).

The analysis reported here is based on the biometric growth-model approach proposed by Neale and McArdle (2000). A schematic representation of the model is given in Fig. 2. Conceptually, the model decomposes an observed phenotypic score (P) into general level (L), and rate-of-linear-change or slope (S) effects and wave-specific or residual effects (E). The L and S effects are general in that they contribute to variance at all four waves of assessment, while the residuals are specific because they contribute to variance at one wave only. Additive genetic, shared environmental, and nonshared environmental factors are allowed to contribute to variance at both the general and wave-specific levels. General effects are correlated, but residual effects are not. The mean scores at the four waves, which for simplicity are not represented in the figure, are modeled in terms of the means for L and S following the logic given by Neale and McArdle (2000). With four waves of assessment, quadratic and cubic as well as linear components of change could in principle be investigated. Because it is very difficult to resolve higher-

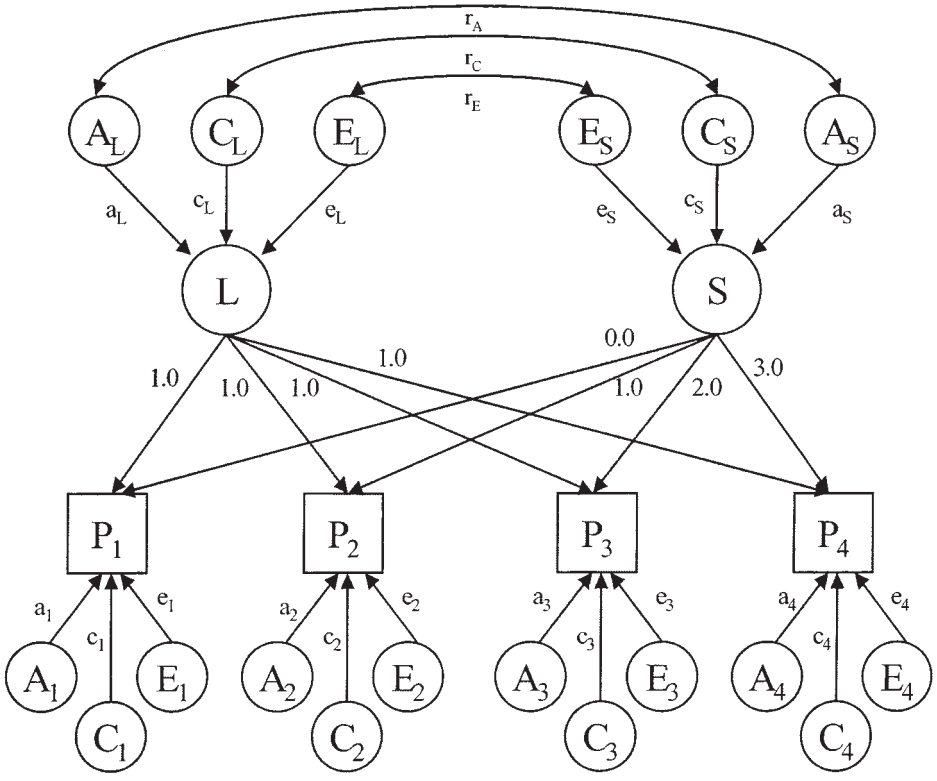

Fig. 2. General linear biometric growth model following Neale and McCardle (2000) fit to the twin data. P_i (i = 1, . . . 4) refers to the observed strength scores at the four assessment waves; L refers to an overall level effect, constrained to contribute equally to variance in the four phenotypes; S refers to a linear rate of change or slope effect; and A_i, C_i, and E_i (i = 1, . . . 4) refer, respectively, to the residual additive genetic, shared environmental, and nonshared environmental effects for the four assessments. The general L and S effects are also decomposed into correlated additive genetic (A_L and As), shared environmental (C_L and Cs), and nonshared environmental (E_L and Es) effects. Lowercase letters refer to the path coefficients and correlations estimated in the growth model. Loadings for L are fixed at 1.0 to ensure it specifies a level effect, and loadings for S are fixed at a linear progression to ensure that it captures the slope effect. Structure for means, although not given in the figure, follows the same structure imposed on the variances.

level polynomial effects with only four waves of data, especially when individuals having three or four observations represent a minority of the sample, however, we have modeled only the linear component of change.

Models were fit to the raw twin data using Mx (Neale *et al.*, 2001), which allowed us to analyze all relevant twin data regardless of the number of assessments completed. In this way we maximized the precision of parameter estimation and the power of hypothesis testing. The raw data option in Mx also allows us to model the missing data (as might occur, for example, due to mortality or participant dropout). Specifically, parameter estimates and model test statistics were adjusted for missing data under the assumption that, conditional on the observations that had been made, the individual's data were missing at random (Little and Rubin, 1987). This treatment of missing data has been shown to produce more precise estimates of parameters than analysis of summary variance-covariance matrices and mean vectors where cases were deleted when even one observation was missing (Greenland and Finkle, 1995).

Models were fit by maximum likelihood by estimating parameters so as to minimize minus two times the natural logarithm of the multivariate normal likelihood (-21nL). The minimized value of this function can also be used to test the relative fit of alternative models. That is, under standard likelihood theory, if -21nL_0 is the minimized value for a base model that involves the estimation of p_0 model parameters, and -21nL_1 is the minimized value for a reduced model that includes a subset of p_1 of these parameters, then the difference in the two minimized functions (i.e., $-21\text{nL}_1 - [-21\text{nL}_0]$) is distributed as a χ^2 statistic on $(p_0 - p_1)$ degrees of freedom (df). We used the differences in -21nL to test a series of null hypotheses following the logic outlined by Neale and McArdle (2000). The Akaike Information Criteria (AIC = $\chi^2 - 2\text{df}$) (Akaike, 1987) was used to supplement the χ^2 goodness-of-fit tests. The AIC balances model fit with model parsimony and is often preferred in large samples like ours where even minor deviations can result in significant χ^2 tests. Models that minimize AIC are preferred.

RESULTS

Twin Resemblance

Twin correlations were estimated in two ways: (1) using only the pairs where both twins participated and had a valid strength score and (2) using simultaneously all available data under the assumption that data were missing at random. Table III shows that the two approaches yielded very similar results except at follow-up 3, where the number of intact pairs is small. The MZ correlation is consistently greater than the DZ correlation except at the last wave, when the sample size drops to only 77 pairs. Because the growth models make use of all available data, they provide a more sensitive test of the existence of genetic influences than the simple comparison of twin correlations embodied in Table III.

Prior to fitting growth models, we fit a series of Cholesky models to determine which components, A, C, and E, were needed in the model and whether there was evidence for sex differences. Table IV gives the fit statistics for the various Cholesky models. AE models fitted better than ACE and CE models, and models that allowed for sex differences fit better than models where parameter estimates were constrained to be equal

Table III. Twin Correlations for Age-Gender Adjusted Strength Scores

	Females		Males	
	MZ	DZ	MZ	DZ
Intake	0.33** (0.33**, $N = 276$)	0.07 (0.06, $N = 411$)	0.11* (0.08, $N = 173$)	0.04 (0.00, $N = 245$)
Follow-up 1	0.31** (0.26**, $N = 155$)	0.11* (0.04, $N = 215$)	0.13* (0.08, $N = 97$)	0.15* (0.16*, $N = 118$)
Follow-up 2	0.31** (0.29**, $N = 66$)	0.15* (0.12, $N = 82$)	0.23* (0.17, $N = 36$)	0.07* (0.06, $N = 40$)
Follow-up 3	0.31* (0.17, $N = 29$)	0.09 (0.14, $N = 24$)	0.04 (-0.11, $N = 11$)	0.22* (0.54*, $N = 13$)

Note: Correlations were estimated under the assumption that missing observations were missing at random and so make use of all available data. Correlations given in parentheses are the traditional intraclass correlations based on only those twin pairs (N) where both participated and had a valid strength score.
* $p < .05$, ** $p < .01$.

Table IV. Cholesky Model Fit Statistics

			Fit relative to general base model			
	−2 ln(L)	df	Δ χ^2	Δ df	*p*	AIC
1. General base model	8075.6	4517	—	—	—	—
Sex differences						
2. Cholesky ACE	8227.5	4617	151.9	100	.001	−48.1
3. Cholesky AE*	8231.8	4637	156.2	120	.015	−83.8
4. Cholesky CE	8241.0	4637	165.4	120	.004	−74.6
5. Cholesky E	8278.5	4657	202.9	140	.0004	−77.1
No sex differences						
6. Cholesky ACE	8279.7	4647	204.1	130	.00004	−55.9
7. Cholesky AE	8279.8	4657	204.2	140	.0003	−75.8
8. Cholesky CE	8289.9	4657	214.3	140	.00005	−65.7
9. Cholesky E	8319.1	4667	243.5	150	< .00001	−56.5

Note: A = additive genetic factor; C = shared environmental factor; E = nonshared environmental factor; AIC = Akaike Information Criterion.
* Best-fitting model according to AIC.

in the male and female samples. Therefore, the growth models were fit separately to the male and female samples and were all based on an AE parametrization.

Table V gives the fit statistics for the growth models that were fit to the twin data. In total there are 16 parameters estimated in the general growth model: (1) 4 parameters corresponding to the general A and E effects for level and slope, (2) 2 parameters for the correlation between the general A and E effects, (3) 8 parameters corresponding to the wave-specific A and E effects, and

Table V. Model Fit Statistics

			Fit relative to model 1 in Table IV			
	−2 ln(L)	df	Δ χ^2	Δ df	*p*	AIC
Full growth model with sex difference	8267.2	4661	191.6	144	.005	−96.4
Orthogonal level and slope						
Males	8269.1	4663	193.5	146	.005	−98.5
Females	8267.6	4663	192.0	146	.006	−100.0
Both	8269.6	4665	194.0	148	.007	−102.0
No general A						
Males	8271.3	4664	195.7	147	.004	−98.3
Females	8300.1	4664	224.5	147	< .0001	−69.5
Both	8304.3	4667	228.7	150	.004	−71.3
No general E						
Males	8568.9	4664	493.3	147	< .0001	199.3
Females	8813.4	4664	737.8	147	< .0001	443.8
Both	9115.1	4667	1039.5	150	< .0001	739.5
No specific A						
Males	8367.2	4665	191.6	148	.009	−104.4
Females	8268.0	4665	192.4	148	.008	−103.6
Both	8268.0	4669	192.4	152	.015	−111.6
No specific A, orthogonal, No sex difference	8317.3	4683	241.7	166	.0001	−90.3
No specific A, orthogonal, sex difference	8270.1	4673	194.5	156	.02	−117.5

Note: Models are based on the path diagram given in Fig. 2. A = additive genetic factor; C = shared environmental factor; E = nonshared environmental factor; AIC = Akaike Information Criterion.

(4) 2 parameters corresponding to the means for the L and S effects. Fit of the growth models was compared with the general base model (i.e., all variances, covariances, and means freely estimated) given in Table IV. As seen in Table V, the analyses of the relative fit of the specific models to the general growth model indicate that the following parameters can individually be fixed to zero without resulting in a statistically significant decrease of model fit: (1) the two general factors correlations ($r_A = 0$, $r_E = 0$) and (2) wave-specific additive genetic effects ($A_i = 0$ for i = 1, . . . 4.). Deleting the general A or E effects on L and S results in a significant increment in the χ^2 test statistic for females, while for males general A was statistically nonsignificant (note that the specific E effects must be included in all models to account for the effects of measurement error). The last model in Table V (no general factor correlations and no specific A) fits the data well relative to the general growth model when sex differences are allowed (χ^2 = 2.9 on 12 df, $p > .99$) and gives minimum AIC (−117.5) among all the models tested.

This model contains 10 estimated parameters for each sex: (1) uncorrelated A effects for level and slope, (2) uncorrelated E effects for level and slope, (3) a specific E effect for each assessment wave, and (4) means for level and slope (Fig. 3).

Parameter estimates for the best-fitting model are given in Table VI. The change in the total variance in strength score over the four waves is statistically significant for both males and females, while the change in genetic variance is statistically significant for females only. There is an approximate doubling of both the total variance and the genetic variance in strength score over the four waves and, hence, a rather stable heritability. However, the heritability in all four waves is around .10 for males and around .30 for females for strength scores. The heritability of level and slope showed a similar pattern: .11–14 in males and .35–39 in females.

DISCUSSION

This study of Danish twins aged 70+ followed longitudinally with assessments every 2 years for up to 6 years showed that there was an overall approximate doubling in variance in physical functioning over 6 years, for both males and females. We also observed an approximate doubling of the genetic variance over this time span. Hence, the heritability was stable over time

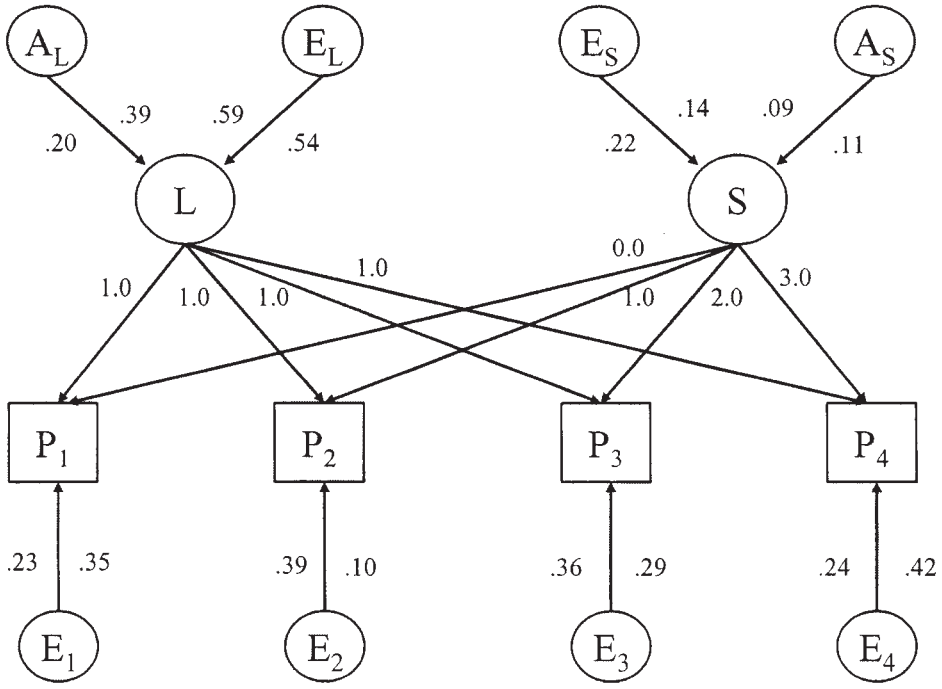

Fig. 3. Best-fitting model was a sex-specific model with no correlation between level (L) and slope (S) effects, no shared environmental effects at either the general or specific levels, and no wave-specific genetic effects (path coefficients (males/females)).

Table VI. Parameter Estimates from Best-Fitting Model

	Intake	Follow-up 1	Follow-up 2	Follow-up 3	Level	Slope
Mean						
Males	3.1	2.9	2.6	2.4	3.1	−.21
Females	3.1	2.9	2.7	2.6	3.1	−.18
Variance total						
Males	.45	.61	.76	.97	.40	.057
Females	.57	.58	.66	.90	.45	.031
Variance genetic						
Males	.04	.05	.07	.11	.04	.008
Females	.16	.17	.21	.27	.16	.012
Variance environmental						
Males	.41	.56	.69	.86	.36	.049
Females	.41	.41	.45	.63	.29	.019
Heritability						
Males	.09	.08	.10	.11	.11	.14
	(.00, .23)	(.00, .19)	(.00, .23)	(.00, .31)	(.00, .26)	(.00, .49)
Females	.28	.29	.31	.30	.35	.39
	(.18, .35)	(.20, .38)	(.21, .41)	(.17, .41)	(.23, .46)	(.03, .72)

Note: Best-fitting model is given in Fig. 3. The change in total variance over the four waves is statistically significant for both males and females, while the change in genetic variance is statistically significant for females only. Numbers given in parentheses are 95% confidence intervals.

but at a considerably higher level for females than for males. This pattern with an increase in both additive genetic variance and environmental variance is in agreement with the genetic-evolutionary and the epidemiological-gerontological theories of aging, respectively.

There are a number of limitations in the study. In relation to the genetic-evolutionary theory it is a central question whether physical abilities at age 70+ can be regarded as a fitness trait. Morphological features are usually not considered to be fitness traits, but physical abilities at older ages correlate to strength measures in midlife (Rantaner *et al.,* 1999), and it seems very likely that strength in midlife in the course of human evolution has been associated with reproductive success and offspring survival. Another concern regarding the present study could be that the follow-up time is limited to 6 years. However, there is a substantial change in both mean and variance over this short period, and the overall change in variance is approximately two fold. Finally, it must be considered to which degree the results are dependent on the growth model implemented. It is reassuring, though, that the censored twin correlations reported in Table III (based simply on the twin pairs in which both participated and had a valid strength score) are very similar to the growth model–derived correlations, except in the fourth wave, where also the numbers of the intact pairs were small.

The analyses of the heritability of the level and the rate of change in this study confirm our analyses of the first three waves, which showed that the level of physical abilities in females is moderately heritable (.3–4), whereas the corresponding heritability in males is nonsignificant (Christensen *et al.,* 2002). While the analyses of the first three waves yielded no evidence for a genetic component to the rate of change, the new finding here is that the four-wave analysis was able to detect a significant heritability of the slope of the same magnitude as on the level. It has recently been suggested that rate of change phenotypes may be especially appropriate targets for molecular genetic investigations aimed at identifying the specific genes influencing human aging (NIA Aging and Genetic Epidemiology Working Group, 2000). This recommendation does not appear to receive support in the previous investigations of rate of change phenotypes that have reported relatively low levels of heritability (Christensen *et al.,* 2002). Even if the present study did detect a moderate heritability for rate of change when all four waves were included, it may still be more promising to focus on mean level of physical functioning at the present early stage of identifying genetic factors of importance for the aging process. The limits of rate of change measures are evident in our findings. First, the assessment of rate of change is especially vulnerable to sample attrition due to mortality and nonresponse; it is not possible to

reliably establish a rate of change phenotype for individuals who do not participate in multiple waves. Second, the rate of change for physical abilities appears to be only moderately heritable, at least over the time span we explored. Everything else being equal, it likely will be easier to identify genes for phenotypes that are more rather than less heritable. Although the estimated heritability of level of physical functioning did not exceed that for rate of change, environmentally induced fluctuations are more easily controlled for the former as compared with the latter by aggregating multiple observations. Nonetheless, our failure to find a significant genetic correlation between the level and slope effects suggests that distinct genetic factors influence the two aspects of late-life physical functioning. Consequently, if the limitations associated with the assessment of rate of change phenotypes could be overcome (e.g., by assessing rate of change over an extended interval from middle to late age), molecular genetic inquiry may be warranted.

We have recently reported on hand grip strength in 1757 Danish twin pairs aged 45–96 (including the twins comprised in the present study, who had grip strength assessed as part of LSADT-1999 and LSADT-2001). Grip strength discriminates functioning in all age-groups, correlates with Activities of Daily Living (ADL) function (Nybo *et al.* 2001), predicts incident disability (Rantanen *et al.*, 1999), and is highly correlated with muscular power in other muscular groups (Rantaner *et al.*, 1994). We found that grip strength had a heritability of 52% (95% confidence interval [CI]: 48% to 55%) in both sexes across all age-groups studied. For males this contrasts with the present study, where we found a small heritability for the strength scale. This is surprising, since the two measurements correlate among LSADT participants (r = .40–.50), and the two-year test-retest correlation for the strength score is approximately .7 and for hand grip .8, indicating reliable measurements. The reason for this discrepancy is unknown. The high heritability found for grip strength across sex and age-groups indicates that grip strength is a suitable phenotype for identifying genetic variants of importance to middle- and late-life physical functioning, although we cannot be sure that the same genetic factors are acting at middle age and among the oldest-old. Currently we are using hand grip level in a study of extreme concordant and extreme discordant DZ pairs from LSADT-1999 and LSADT-2001 in a genome-wide search for genetic factors affecting physical functioning in second half of life (Frederiksen *et al.*, 2002).

ACKNOWLEDGMENTS

This work was supported by the U.S. National Institute on Aging research grant NIA-P01-AG08761 and the Danish National Research Foundation.

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. *Psychometrica* **52**:317–332.

Avlund, K., Davidsen, M., and Schultz-Larsen, K. (1995). Changes in functional ability from age 70 to age 75: A Danish longitudinal study. *J. Aging Health* **7**:254–282.

Bryk, A. S., and Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychol. Bull.* **101**:147–158.

Campisi, J. (2001). From cells to organisms: Can we learn about aging from cells in culture? *Exp. Gerontol.* **36**:607–618.

Charlewood, B., and Hughes K. A. (1996). Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence. *Proc. Natl. Acad. Sci. USA* **93**:6140–6145.

Christensen, K., Gaist, D., Vaupel, J. W., and McGue, M. (2002). Genetic contribution to rate of change in functional abilities among Danish twins aged 75 years and older. *Am. J. Epidemiol.* **155**:132–139.

Christensen, K., Holm, N. V., McGue, M., Corder, L., and Vaupel, J. W. (1999). A Danish population-based twin study on general health in the elderly. *J. Aging Health* **11**:49–64.

Christensen, K., McGue, M., Yashin, A. I., Iachine, I. A., Holm, N. V., and Vaupel, J. W. (2000). Genetic and environmental influences on functional abilities among Danish twins aged 75 years and older. *J. Gerontol. Med. Sci.* **55A**:M446–M452.

Frederiksen, H., Gaist, D., Petersen, H. C., Hjelmborg, J., McGue, M., Vaupel, J. W., and Christensen, K. (2002). Hand grip strength: A phenotype suitable for identifying genetic variants affecting mid and late life physical functioning. *Genet. Epidemiol.* **23**:110–122.

Greenland, S., and Finkle, W. D. (1995). A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am. J. Epidemiol.* **142**:1255–1264.

Harris, J. R., Pedersen, N. L., McClearn, G. E., Plomin, R., and Nesselroade, J. R. (1992). Age differences in genetic and environmental influences for health from the Swedish Adoption/Twin Study of Aging. *J. Gerontol. Psychol. Sci.* **47**:P213–P220.

Hauge, M. (1981). The Danish Twin Register. In S. A. Mednick, A. E. Baert, and B. Bachmann (eds.), *Prospective longitudinal research* (pp. 217–222). Oxford: Oxford Medical Publications.

Holm, N. V. (1983). The use of twin studies to investigate causes of diseases with complex etiology with focus on cancer. Odense University, Odense, Denmark.

Hughes, K. A., and Burleson, M. H. (2000). Evolutionary causes of genetic variation in fertility and other fitness components. In J. L. Rodgers, D. C. Rowe, and W. B. Miller (eds.), *Genetic influence on human fertility and sexuality: Theoretical and empirical contributions from the biological and behavioral sciences* (pp. 7–33). Boston: Kluwer Academic Publishers.

Jette, A. M. (1987). The functional status index: Reliability and validity of a self-report functional disability measure. *J. Rheumatol.* **14**:(Suppl. 15):15–19.

Kane, R. A., and Kane, R. L. (1981). Assessing the elderly: A practical guide to measurement. Lexington, MA: Lexington Books.

Katz, S, Ford, A. B., Moskowitz, R. W., *et al.* (1963). Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *JAMA* **185**:94–112.

Kjøller, M. (1995). *Health and morbidity in Denmark*. Copenhagen: Danish Institute for Clinical Epidemiology.

Little, R. T. A., and Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Manton, K. G., Corder, L., and Stallard, E. (1993). Changes in the use of personal assistance and special equipment from 1982 to 1989: Results from the 1982 and 1989 NLTCS. *Gerontologist* **33**:168–176.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., and Woodcock, R. W. (2002). Comparative longitudinal analyses of the growth and decline of multiple intellectual abilities over the life span. *Dev. Psychol.* **38**:115–142.

McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., and Plomin, R. (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science* **276**(5318):1560–1563.

McGue, M., and Bouchard, T. J. J. (1984). Adjustment of twin data for the effects of age and sex. *Behav. Genet.* **14**:325–343.

McGue, M., and Christensen, K. (2001). The heritability of cognitive functioning in very old adults: Evidence from Danish twins aged 75 and older. *Psychol. Aging* **16**:272–280.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (2001). *Mx: Statistical modeling* (5th ed.), Box 126 MCV, Richmond VA 23298: Department of Psychiatry.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic.

Neale, M. C., and McArdle, J. J. (2000). Structure latent growth curves for twin data. *Twin Res.* **3**:165–177.

NIA Aging and Genetic Epidemiology Working Group. (2000). Genetic epidemiologic studies on age-specified traits. *Am. J. Epidemiol.* **152**:1003–1008.

Nybo H., Gaist D., Jeune B., McGue M., Vaupel J. W., Christensen K. (2001) Functional status and self-rated health in 2,262 nonagenarians: the Danish 1905 Cohort Survey. *Journal of the American Geriatrics Society*: **49**:601–609.

Patridge, L. (2001). Evolutionary theories of ageing applied to long-lived organisms. *Exp. Gerontol.* **36**:641–650.

Platz, M. (1989). *The elderly in their homes*: Vol. 1. *Living conditions*. Copenhagen: Danish National Institute of Social Research.

Platz, M. (1990). *The elderly in their homes*: Vol. 2. *How do they cope?* Copenhagen: Danish National Institute of Social Research.

Plomin, R., DeFries, J. C., McClear, G. E., and McGuffin, P. (2001). *Behavioral genetics* (4th ed.). New York: Freeman.

Rantaner, T., Guralnik, J. M., Foley, D., Masaki, K., Leveille, S., Curb, J. D., and White, L. (1999). Midlife hand grip strength as a predictor of old age disability. *JAMA* **281**:558–560.

Rantaner, T., Pertti, E., Kauppinen, M., and Heikkinen, E. (1994). Maximal isometric muscle strength and socioeconomic status, health, and physical activity in 75-year-old persons. *J. Aging Physical Activity* **2**:206–220.

Schulz-Larsen, K., Avlund, K., and Kreiner, S. (1992). Functional ability of community-dwelling elderly: Criterion-related validity of new measure of functional ability. *J. Clin. Epidemiol.* **45**:1315–1326.