Rates of Aging, *APOE* Genotypes, and Cause-Specific Mortality in the Cache County Study

Ken R. Smith
University of Utah

Kathleen Hayden
Duke University

Geri Mineau
University of Utah

Richard Kerber
University of Utah

September 22, 2005

Extended Abstract

Introduction: Assessments of mortality risks among the elderly have had a long history in demography. These studies have generally considered nominal biological measures such as gender and age. Recent decades have seen a growing list of risk factors included in such analyses such as serum cholesterol, blood pressure, and body mass index. With the expansion of molecular biology and its arsenal of techniques for identifying genetic markers for disease and wellness susceptibility, there is now an increasing availability of new genotype information on large population-based samples. This development now affords demographers with an excellent opportunity to integrate these measures into analyses of mortality differentials.

Objectives: The purpose of this study is to assess the role of the Apolipoprotein E (*APOE*) gene in affecting cause-specific mortality risks in a population-based cohort of elderly subjects. The study relies on 4701 subjects recruited into the Cache County Study on Memory and Aging. The association between *APOE* genotypes and all-cause and cause-specific mortality has been

examined by numerous investigators [1, 2]. A recent study based on the Cache County study showed that the risk of all-cause mortality persons doubled for persons with ϵ4/ϵ4 in relation to those with ϵ3/ϵ3 [3]. A key feature of this growing literature is the heterogeneity of the mortality risk estimates associated with *APOE* genotypes. These discrepant results may be explained by the variety of samples and methods used. Specifically, less attention has been given to potential confounders and effect modifiers of the mortality effects of *APOE* alleles, except for the possible effects of diet [4, 5]. We suggest that much can be learned about the varying mortality effects of *APOE* by focusing on its interactions with measures of rates of aging. The biodemographic literature on longevity has revitalized interest in the family history of longevity and female reproductive senescence as key concepts in understanding aging and mortality in the later years and individual rates of aging [6-10]. An examination of the joint influences of a known genetic factor associated with later life mortality (*APOE*) and biodemographic measures of rates of aging is a useful strategy for exploring the heterogeneous effects of *APOE* genotypes on later-life mortality.

In summary, this paper seeks to accomplish the following:

1. Generate estimates of the relative risk of *APOE* ϵ2/ϵ2, ϵ2/ϵ3, ϵ2/ϵ4, ϵ3/ϵ4 and ϵ4/ϵ4, in relation to the *APOE* ϵ3/ϵ3 on the risk of death from hart disease, cancer, respiratory diseases, and diseases of the nervous system.
2. Estimate how rates of aging, as measured by familial excess longevity and female reproductive aging (from UPDB), affect the risk of mortality from risk of death from hart disease, cancer, respiratory diseases, and diseases of the nervous system.

3. Estimate how the association between *APOE* genotype and cause-specific mortality is modified by indicators of rates of aging.

Study Design: This analysis relies on the Cache County Study on Memory and Aging, a longitudinal study of dementia and cognitive health based on a large representative sample of all older adult residents (aged 65+) of Cache County, Utah. Briefly, all residents of the county aged 65+ as of 1 January 1995 were invited to participate. Procedures were approved by the institutional review boards of Utah State University, Johns Hopkins University, and Duke University. Participants were asked to provide buccal DNA for genotyping at the *APOE* locus. They were then asked to complete a two-hour interview on health factors potentially related to dementia, including cardiovascular disorders, and they were evaluated for dementia using a multi-stage assessment procedure (Wave 1). Informed consent was obtained when participants were able to understand the procedure, and proxy consent was obtained from spouses or next of kin for others. Approximately three years after the baseline evaluation, between 1998 and 1999, survivors without dementia at Wave 1 were asked to complete a follow-up interview to assess changes in health factors, and were again evaluated with a similar multi-stage procedure to detect new cases of dementia (Wave 2). Mortality among the participants was monitored continuously through 25 January 2002.

Study Sample: Of 5,956 eligible individuals 5,092 completed the baseline interview. Most (4,969) were also genotyped at the *APOE* locus. Of these, 261 refused to complete the Wave 1 evaluations, often because of medical illness. The remaining 4,708 participants were included in these analyses with 4650 having complete genotype information.

The Cache County Study on Memory and Aging has recently been enriched with the addition of several important variables obtained from the Utah Population Database. These variables include measures of familial excess longevity (FEL as defined below) along with cause of death information obtained from death certificates of Cache County Study decedents.

The Utah Population Database (UPDB) contains over seven million records, including the genealogies of the founders of Utah and their descendants. In the 1970s, approximately 170,000 Utah nuclear families were identified on "Family Group Sheets" from the archives at the Utah Family History Library, each with at least one member having had a vital event (birth, marriage, death) on the Mormon Pioneer Trail or in Utah. These families have been linked across generations; in some instances, the records span seven generations. The UPDB is an active genealogy; new families and their members are continually being added as the UPDB is linked to other sources of data, including birth and death certificates. Additional information on these individuals comes from sources such as drivers' license and the Utah Cancer Registry. The UPDB now holds data from migrants to Utah and their descendants that number more than 1.8 million individuals born from the early 1800's to the mid-1900's and that are linked into multi-generation pedigrees. Through funding from the NIA, nearly all Cache County study subjects have been linked to the UPDB.

We use familial excess longevity (FEL) as a genealogical-based method for measuring slow rates of aging. FEL is a summary measure of excess longevity among all blood relations for a given individual. Calculating FEL first requires that an estimate of the difference between an

individual's attained age and the age to which that individual was expected to live according to an accelerated failure time model. This model uses three covariates that are available in the UPDB and that are associated with longevity: gender, birth year, affiliation with the Church of Jesus Christ of Latter-day Saints (or LDS Church or Mormons). The variable describing whether a person in the UPDB is a member of the LDS Church is included because it is known that the church members have significantly longer lives.

The expected age at death is:

$$\hat{y} = \exp(\beta_0 + \beta_1\, Gender + \beta_2\, BirthYear + \beta_3\, Religious\,Affiliation),$$

with β_i, i = 0, . . . , 3, regression coefficients. The excess longevity is $l = y - \hat{y}$, where y is the age at death or the age at the time last confirmed alive, in years. We focus on only those subjects who reached the age of 65. The familial excess longevity for a subject is calculated as the weighted average of individual excess longevity of all blood relatives. The weights are the kinship coefficients, the probability, K_{ij}, that individual i shares a particular allele with individual j. The familial excess longevity for subject i is

$$FEL_i = \frac{\sum_{j \in J} K_{ij} l_j}{\sum_{j \in J} K_{ij}}$$

where J is the set of all blood relatives of subject i living to age 65.

In summary, the merged Cache County-UPDB data set offers an excellent opportunity for studying aging and risk factors for cause-specific mortality past age 65 with *APOE* genotype information, a genealogically based measure of rates of aging (FEL), cause of death information on all decedents, and a vast set of key covariates measuring known risk factors for mortality

among the elderly.

Statistical Methods: The outcome of interest is age at death and cause of death. Given that the analysis examines survival, we use Cox proportional hazards regression models and its extensions. Each cause-of-death of interest will be estimated using standard competing risk survival techniques (i.e., each cause of death treats the remaining causes of death as uninformative censored cases). The basic model will include covariates that measure *APOE* genotype, familial longevity, and reproductive senescence along with their interaction terms. The other variables requested are to adjust for confounding and to help explicate the pathways by which *APOE* genotypes affect cause-specific mortality.

***Results*:** In Table 1, the descriptive statistics for the key covariates are shown. Table 2 provides the distribution of the causes of death for the cohort. Note that death from the circulatory system, neoplasms, respiratory diseases, and diseases of the nervous system comprise the largest categories of decedents and are therefore the focus of the cause-specific mortality analysis. Those dying from neoplasms died primarily from (female) breast, prostate, colon, and lung cancer. Respiratory disease deaths were mostly from pneumonia while deaths from diseases of the nervous system were from Parkinson's disease and Alzheimer's disease.

Tables 3 through 7 provide regression coefficient estimates for each of the *APOE* genotypes (in relation to $\epsilon 3/\epsilon 3$) based on Cox proportional hazard models. These first set of models focus on objective 1 at this point while controlling for the effects of gender, age, and education.

In making an assessment of each of the *APOE* genotypes, we find that ϵ4/ϵ4 significantly increases the risk of all-cause mortality as well as for mortality from heart disease, respiratory disease, and diseases of the nervous system. We also find that ϵ2/ϵ2 significantly increases the risk of mortality from neoplasms and diseases of the nervous system.

Table 8 and 9 show results of the effects of *APOE* genotypes, FEL, and their interaction on all-cause mortality. Table 8 includes the main effects of FEL without interaction. Note that for some subjects FEL was not estimable. In these instances, the mean value of FEL was substituted for the missing FEL cases and a dummy variable, FELMISS, was included to identify those whose value of FEL was imputed with the sample mean. This procedure preserved the full sample. The results in Table 8 for FEL show that those with increasing values of FEL had significantly lower mortality rates. In Table 9, where *APOE* genotypes, FEL, and their interactions are included, we find that the adverse effects of mortality associated with ϵ3/ϵ4 and ϵ4/ϵ4 are significantly attenuated for persons with increasing FEL.

***Summary*:** These preliminary results indicate how our understanding of aging and mortality risks can be affected by APOE genotypes. We show that the deleterious effects of APOE genotypes (e22, e34, e34) occur for several important cause-specific risks of mortality (cardiovascular, cancer, respiratory, nervous system (Parkinson's, Alzheimer's) and that their effects are attenuated when they occur to individuals with a family history of excess longevity. The strategy of using genotype information along with powerful indices of rates-of-aging can be used to contribute to our understanding of familial and genetic forces affecting the longevity of the elderly.

References

1. Eichner, J.E., et al., *Apolipoprotein E polymorphism and cardiovascular disease: A HuGE review.* American Journal of Epidemiology, 2002. **155**(6): p. 487-495.
2. Lee, J.H., et al., *Mortality and apolipoprotein E in Hispanic, African-American, and Caucasian elders.* American Journal of Medical Genetics, 2001. **103**(2): p. 121-127.
3. Hayden, K., et al., *APOE genotype and mortality: The Cache County study.* Gerontologist, 2003. **43**: p. 8-8.
4. Ordovas, J.M., *Gene-diet interaction and plasma lipid responses to dietary intervention.* Biochemical Society Transactions, 2002. **30**(2): p. 68-73.
5. Schaefer, E.J., *Lipoproteins, nutrition, and heart disease.* The American Journal of Clinical Nutrition, 2002. **75**(2): p. 191-212.
6. Kerber, R.A., et al., *Familial excess longevity in Utah genealogies.* J Gerontol A Biol Sci Med Sci, 2001. **56**(3): p. B130-9.
7. Smith, K.R., G.P. Mineau, and L.L. Bean, *Fertility and post-reproductive longevity.* Social Biology, 2002. **49**(3-4): p. 185-205.
8. Vaupel, J.W., et al., *Biodemographic trajectories of longevity.* Science, 1998. **280**(5365): p. 855-60.
9. Wachter, K.W., C.E. Finch, and National Research Council (U.S.). Committee on Population., *Between Zeus and the salmon : the biodemography of longevity*. 1997, Washington, D.C.: National Academy Press. viii, 285.
10. Yashin, A.I., et al., *Genes, demography, and life span: the contribution of demographic data in genetic studies on aging and longevity.* Am J Hum Genet, 1999. **65**(4): p. 1178-93.
11. Lehrer, S., *Possible relationship of the apolipoprotein E (ApoE) epsilon 4 allele to prostate cancer.* British Journal of Cancer, 1998. **78**(10): p. 1398-1398.
12. Moysich, K.B., et al., *Apolipoprotein E genetic polymorphism, serum lipoproteins, and breast cancer risk.* Molecular Carcinogenesis, 2000. **27**(1): p. 2-9.
13. Watson, M.A., et al., *Apolipoprotein E gene polymorphism and colorectal cancer: gender-specific modulation of risk and prognosis.* Clinical Science, 2003. **104**(5): p. 537-545.

Table 1. Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
age	4708	76.112	7.218	64.701	105.742
age2	4708	5845.070	1133.900	4186.270	11181.470
EDUC	4703	13.199	2.869	0.000	20.000
male	4708	0.435	0.496	0.000	1.000
APOE					
ϵ2/ϵ2	4650	0.007	0.084	0.000	1.000
ϵ2/ϵ3	4650	0.132	0.339	0.000	1.000
ϵ2/ϵ4	4650	0.036	0.186	0.000	1.000
ϵ3/ϵ3	4650	0.543	0.498	0.000	1.000
ϵ3/ϵ4	4650	0.256	0.437	0.000	1.000
ϵ4/ϵ4	4650	0.026	0.159	0.000	1.000
FEL	4708	3.411	1.736	-10.230	16.720
duration2002	4708	1872.970	768.990	4.000	2919.000
censor2002	4708	0.655	0.475	0.000	1.000

Table 2. Frequency Distribution for Causes of Death

Cause of Death	Frequency	Percent
Alive	3085	65.53
Circulatory System	633	13.45
Neoplasms	249	5.29
Dead but not classified	181	3.84
Respiratory	176	3.74
Nervous System/Sense Organs	84	1.78
Endocrine/Metabolic	62	1.32
Digestive System	57	1.21
Mental/Psychoneurotic	42	0.89
External Causes	35	0.74
Symptoms/Senility/Ill-Defined	31	0.66
Genito-urinary	30	0.64
Bones/Organs of Movement	17	0.36
Infection/Parasite	13	0.28
Skin/Subcutaneous Tissue	6	0.13
Blood/Blood-forming Organs	4	0.08
Congenital Malformations	3	0.06

Table 3. Effects of APOE Genotypes on All-Cause Mortality

Analysis of Maximum Likelihood Estimates				
	Parameter Estimate	**Standard Error**	**P**	**Hazard Ratio**
age	0.27381	0.06284	<.0001	1.315
age2	-0.00106	0.0003879	0.0063	0.999
EDUC	-0.02387	0.00874	0.0063	0.976
male	0.28642	0.05076	<.0001	1.332
ϵ2/ϵ2	0.53327	0.25259	0.0348	1.704
ϵ2/ϵ3	-0.18501	0.08057	0.0217	0.831
ϵ2/ϵ4	-0.22746	0.15133	0.1328	0.797
ϵ3/ϵ4	0.12486	0.0599	0.0371	1.133
ϵ4/ϵ4	0.28844	0.1602	0.0718	1.334

Table 4. Effects of APOE Genotypes on Cardiovascular Mortality

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.22564	0.08668	0.0092	1.253
age2	-0.0007742	0.0005349	0.1478	0.999
EDUC	-0.0353	0.01245	0.0046	0.965
male	0.23538	0.07167	0.001	1.265
ϵ2/ϵ2	0.37318	0.38134	0.3278	1.452
ϵ2/ϵ3	-0.19919	0.11346	0.0792	0.819
ϵ2/ϵ4	-0.33926	0.22335	0.1288	0.712
ϵ3/ϵ4	0.0703	0.08549	0.4109	1.073
ϵ4/ϵ4	0.42595	0.21015	0.0427	1.531

Table 5. Effects of APOE Genotypes on Cancer Mortality

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.49009	0.13806	0.0004	1.632
age2	-0.00271	0.0008712	0.0019	0.997
EDUC	-0.02818	0.01732	0.1036	0.972
male	0.29486	0.09837	0.0027	1.343
ϵ2/ϵ2	0.83634	0.38381	0.0293	2.308
ϵ2/ϵ3	-0.18285	0.15281	0.2315	0.833
ϵ2/ϵ4	-0.49948	0.32277	0.1217	0.607
ϵ3/ϵ4	-0.09847	0.11998	0.4118	0.906
ϵ4/ϵ4	0.21799	0.2962	0.4618	1.244

Table 6. Effects of APOE Genotypes on Respiratory Disease Mortality

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.17274	0.12934	0.1817	1.189
age2	-0.0004079	0.0007964	0.6085	1
EDUC	-0.04253	0.01888	0.0243	0.958
male	0.21945	0.1086	0.0433	1.245
ϵ2/ϵ2	0.8769	0.45372	0.0533	2.403
ϵ2/ϵ3	-0.19794	0.16973	0.2435	0.82
ϵ2/ϵ4	0.11615	0.26785	0.6646	1.123
ϵ3/ϵ4	-0.14503	0.1388	0.2961	0.865
ϵ4/ϵ4	0.63941	0.28731	0.026	1.895

Largely pneumonia

Table 7. Effects of APOE Genotypes on Nervous System Mortality

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.11122	0.15193	0.4642	1.118
age2	-0.0001541	0.0009458	0.8705	1
EDUC	-0.04463	0.02206	0.0431	0.956
male	0.199	0.12526	0.1121	1.22
ϵ2/ϵ2	0.23037	0.71279	0.7466	1.259
ϵ2/ϵ3	-0.2899	0.21008	0.1676	0.748
ϵ2/ϵ4	0.14004	0.31318	0.6548	1.15
ϵ3/ϵ4	0.14762	0.14764	0.3174	1.159
ϵ4/ϵ4	0.76218	0.31416	0.0153	2.143

Largely Parkinson’s and Alzheimer's

Table 8. Effects of APOE Genotypes And Familial Excess Longevity on Al-Cause Mortality - Main Effects only

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.19611	0.06508	0.0026	1.217
age2	-0.000543	0.0004033	0.1782	0.999
EDUC	-0.01791	0.00933	0.0548	0.982
male	0.36974	0.0545	<.0001	1.447
ϵ2/ϵ2	0.67046	0.27014	0.0131	1.955
ϵ2/ϵ3	-0.14746	0.08677	0.0892	0.863
ϵ2/ϵ4	-0.36559	0.17809	0.0401	0.694
ϵ3/ϵ4	0.1752	0.0638	0.006	1.191
ϵ4/ϵ4	0.26037	0.17119	0.1283	1.297
FEL	-0.03738	0.01475	0.0113	0.963
felmiss	0.129	0.0583	0.0269	1.138

Table 9. Effects of APOE Genotypes And Familial Excess Longevity on Al-Cause Mortality - Main and Interaction Effects

Analysis of Maximum Likelihood Estimates				
Variable	**Parameter Estimate**	**Standard Error**	**P**	**Hazard Rate**
age	0.18794	0.06522	0.004	1.207
age2	-0.000491	0.0004044	0.2247	1
EDUC	-0.01723	0.00934	0.065	0.983
male	0.37102	0.05464	<.0001	1.449
ϵ2/ϵ2	0.68395	0.27349	0.0124	1.982
ϵ2/ϵ3	-0.14415	0.08693	0.0973	0.866
ϵ2/ϵ4	-0.37668	0.18108	0.0375	0.686
ϵ3/ϵ4	0.16752	0.06424	0.0091	1.182
ϵ4/ϵ4	0.20041	0.17886	0.2625	1.222
FEL	-0.03709	0.0147	0.0116	0.964
felmiss	0.13523	0.05843	0.0206	1.145
FEL * ϵ2/ϵ2	-0.07902	0.18543	0.67	0.924
FEL * ϵ2/ϵ3	-0.05052	0.0465	0.2773	0.951
FEL * ϵ2/ϵ4	-0.06951	0.07529	0.3559	0.933
FEL * ϵ3/ϵ4	-0.0845	0.03548	0.0173	0.919
FEL * ϵ4/ϵ4	-0.15094	0.06528	0.0208	0.86