Featured Article

Age stratification corrects bias in estimated hazard of *APOE* genotype for Alzheimer's disease

Li Liu[a,b,c,**], Richard J. Caselli[c,*]

[a]*Department of Biomedical Informatics, Arizona State University, Scottsdale, AZ, USA*
[b]*Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, AZ, USA*
[c]*Department of Neurology, Mayo Clinic Arizona, Scottsdale, AZ, USA*

Abstract

Introduction: The apolipoprotein E (APOE) e4 allele is a major genetic risk factor of late-onset Alzheimer's disease. However, its interaction with two other canonical risk factors, age and sex, is not clear. Previous studies have reported conflicting results on its differential effects in men and women, its association with young-onset AD before the age of 65 years, and its significance in genetic admixture populations. In these studies, the hazard of the e4 allele was assumed to be constant during aging. However, this hypothesis has not been tested and its violation may lead to significant biases and contribute to such discrepant findings.

Methods: In a prospective cohort of 4727 subjects, we performed Cox regression analysis of the association of the e4 allele with AD age of onset. We then performed diagnostics on the resulting model and tested if the hazard of the e4 allele violated the assumption of proportionality during aging. We examined whether incorporating age stratifications and time-dependent coefficients could restore the proportionality. We then validated our findings in four independent cohorts.

Results: Hazard of the e4 allele for AD was nonproportional. It took a stepwise decline around the age of 80 years for men and around the age of 75 years for women. By stratifying subjects into a younger group and an older group, we detected more consistent effects of the e4 allele across multiple independent cohorts. We also found that the e4 allele was a significant risk factor for young-onset AD with age of onset before 65 years.

Discussion: Age compositions of study cohorts can significantly bias the estimated effect of the APOE genotype. Studies of AD should consider hidden age structures among subjects and routinely employ appropriate age and sex stratification strategies or nonparametric modeling in experimental designs and data analysis. Finally, we argue that the e4 allele is a risk factor not only for late-onset AD but also for young-onset AD.

© 2018 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Alzheimer's disease; APOE genotype; Disease risk; Age dependency; Young onset dementia

1. Introduction

Age, sex, and apolipoprotein E (APOE) genotype are three major risk factors for Alzheimer's disease (AD) that are known to interact with each other [1–3]. Some effects are well

established, for example, the modification of age of onset by the APOE genotype [4–6]. Other effects are less clear. All studies agree that the APOE e4 allele is a major risk factor for late-onset AD (LOAD), but Farrer *et al.* and van Dujin *et al.* independently reported that the gene dose-effect of the APOE e4 allele was evident as early as 40 years old (yo), whereas others argued that the e4 allele has no significant association with the onset of AD before the age of 65 years [7–9]. Studies examining sex-related predisposition to AD reported

contradictory findings that vary from higher risk to lower risk in women to no difference between the sexes [1,2,10]. There are even published studies that failed to detect any association between AD risk and APOE genotype [11,12].

Of the many factors that may contribute to such discrepancies, for example, ethnicity and cohort size [13–15], age is a common confounding factor. Appropriate manipulation of age-dependent variables is critical to avoid biases. For example, Cox regression is frequently used to estimate the hazard ratio (HR) of the APOE genotype on age of AD onset. Cox regression assumes the risk associated with a specific allele is constant across all ages (i.e., a proportional hazard). When this assumption is violated, it computes an average HR over all ages [16]. Another common method is logistic regression that allows age to be incorporated as a covariate and computes an age-adjusted odds ratio, which is essentially a weighted average over all ages. Thus, given a cohort, the age composition can unexpectedly bias the estimated effect of the APOE genotype in the presence of nonproportional hazards.

To our knowledge, no investigations have yet tested the hypothesis of proportional hazard of the APOE genotype during aging. Although several studies have attempted to group subjects on baseline age or include higher-order terms or interaction terms in a logistic regression model [2,3,8,17], it remains questionable if these efforts were sufficient in modeling the potentially complex time-varying effects. Furthermore, age, sex, and APOE genotype are common covariates when evaluating novel biomarkers and clinical trials. A formal testing of the proportionality and delineation of their relationships will provide valuable guidance to experimental design, model selection, and data interpretation.

In this study, we performed systematic evaluation of the age-dependent effects of the APOE e4 allele on AD risk stratified on sex. We used the uniform data set from the Alzheimer's Disease Centers (ADCs) that has been curated by the National Alzheimer's Coordinating Committee to ensure diagnostic consistency [18]. We reconfirm that APOE e4 is a powerful risk factor for LOAD and also show that it is a significant risk factor for young-onset AD (YOAD). We demonstrated that the assumption of proportionality was violated in both men and women. We then identified critical change points that marked the transition of a stepwise decline of the hazards of the e4 allele during aging. Informed by these change points, we conducted age stratifications in four independent cohorts and rectified the biases of the estimated hazards of the e4 allele. We further explored an alternative stratification strategy using Bayesian analysis to account for the nonproportionality in model fitting. We discussed the implications of these discoveries for future AD studies.

2. Methods

We downloaded the ADC uniform data set [18] from the National Alzheimer's Coordinating Committee database (last accessed on 12/12/2016) that included longitudinal observations of 11,057 subjects. By removing subjects with prevalent dementia, dementia with non-AD causes, or non-Caucasian ancestry, we obtained 3731 subjects with normal cognitive function at the last follow-up and 996 subjects with incident AD. For clinically unaffected controls, we used the age at last follow-up as the age at significant event. For patients with AD, we used the age of onset as the age at significant event. To account for the delay between disease onset and diagnosis, we took the age at diagnosis and the age at the immediately preceding epoch with normal cognitive function and used the average age between those as age of onset. Because e2 allele counts were reported for only 9% (407) of these subjects, we restricted our analysis to e4 allele counts.

Using the ADC data, we conducted a standard multivariate Cox regression analysis to test if *e4* allele count, sex, and their pairwise interactions were significantly (two-tailed *P* value < .05) associated with the age of AD onset. We then analyzed the data for men and women separately. For each Cox regression model, we performed diagnostics by computing the scaled Schoenfeld residuals (R_{ss}) and testing if the residuals varied with age of onset using Pearson's test of goodness-of-fit [19]. *P* value < .05 indicated a violation of the proportionality assumption. For graphical diagnostics, we plotted a smoothed curve of R_{ss} against age of onset (residual-time curve). This curve illustrates if the time-dependent HR deviates from a horizontal line (i.e., slope = 0 for time-independence) [19]. When the diagnostic statistics indicated time-varying HRs, we searched the critical change point in a grid between ages 55 and 95 years with 0.5 year as the interval. For each candidate change point (*T*), we extended the Cox regression model to include a time-dependent covariate *z* and defined the hazard rate *h* at age *t* as

$$h(t|x,\ z)=h_0(t)\exp(\beta_1 x+\beta_2 z) \qquad (1)$$

where *x* was the e4 allele count, h_0 (*t*) was the baseline hazard rate, *z* was a time-dependent covariate (*z* = 1 if *t* > *T*, *z* = 0 otherwise), and β_1 and β_2 were the coefficients to estimate. We then computed the log partial likelihood of the resulting model. The value of *T* giving the highest partial likelihood was chosen as the transition age to stratify the subjects. Note that equation (1) enables the realization of time-varying coefficients. The relative risk of an e4 allele is exp (β_1) before age *T* and is exp (β_1 + β_2) after age *T*. Given the critical change point, we used this extended Cox regression to estimate an HR for each age stratum. To compare two HRs, we conducted a two-sided z-test. *P* value < .05 indicated that two HRs were significantly different. We used the R/survival package for Cox regression analysis and diagnostics [20].

To validate our findings, we used the discovery set of the Alzheimer's Disease Sequencing Project (dbGaP study phs000572.v7.p4, last accessed on 02/15/2017). After

removing subjects with prevalent AD, undetermined AD status, or non-Caucasian ancestry, we found four independent cohorts (Adult Changes in Thought, Cardiovascular Health Study, Framingham Heart Study, and Rotterdam Study) that had no overlap with the ADC uniform data and at least 300 subjects. In total, these four cohorts contained 1972 clinically unaffected individuals and 756 incident AD patients. We applied the transition age identified in the ADC uniform data to each of the four cohorts and computed HRs of e4 allele counts for each age group.

As an alternative to age stratification, we examined stratifications on hazards using the Bayesian analysis for a linear dependent Dirichlet process (LDDP) mixture of survival models that is implemented in the R/DPpackage [21]. We carried out inferences with e4 allele counts as the covariate for men and for women separately. Specifically, we called the LDDPsurvival function with priors set to a0 = 10, b0 = 1, nu = 4, m0 = c(2,2), S0 = diag(100,2), psiinv = diag(1,2), tau1 = 2.01, taus1 = 2.01, and taus2 = 1.01. We used 1000 burn-ins with a thinning interval of 5 during the Markov Chain Monte Carlo (MCMC) sampling. We also varied the priors to examine the influence on clustering patterns. Given the ADC data, we computed the posterior estimate of pointwise hazard ratio as

$$E(HR \mid data, \ t) = \log(S_1(t))/\log(S_0(t)) \qquad (2)$$

where $S_0(t)$ and $S_1(t)$ are the posterior probability of survival at time t for e4 allele count = 0 and 1, respectively.

3. Results

Subjects involved in this analysis had two to eleven visits to ADCs between 2005 and 2016 for cognitive evaluations. The average interval between visits for each subject was 11 months (range: 4 months–3 years). 62% of the subjects were women. The percentages of e4 allele carriers in men and in women were similar (0.33 vs. 0.31, Fisher's exact test *P* value = .07). Additional characteristics of subjects are available in Table 1. Standard Cox regression analyses with sex and e4 allele counts as covariates found that these two factors had significant interactions (*P* value = .02). Thus, we stratified subjects by sex and built a univariate Cox regression model for e4 allele counts in men and women separately.

These standard univariate Cox regression models assumed proportional hazards. We performed diagnostics on each model to examine if the effects of e4 allele counts varied with age. Goodness-of-fit tests on R_{ss} over time showed that the HRs of e4 allele counts in men and for women both deviated significantly from proportionality (*P* value = 3×10^{-4} and 5×10^{-4}, respectively). The residual-time curves had a stepwise decline, showing higher HRs in younger ages than in older ages (Fig. 1A, Supplementary Tables 2-4). A grid search of critical change points of HRs found that 80.5 yo and 74.5 yo were the transition age for men and women, respectively

Table 1
Characteristics of the ADC and ADSP cohorts (ACT, CHS, FHS, and RS)

Cohort	ADC	ACT	CHS	FHS	RS
Total Subjects	4727	1134	612	350	632
Sex					
Male	1787	500	258	154	211
Female	2940	634	354	196	421
Subjects without AD	3731	830	439	249	454
Age at last follow-up, years					
Mean	77.4	86.5	84.1	84.7	85.9
Min	53	68	76	61	75
Max	104	89	89	90	90
APOE e4 count					
0	2742	707	398	216	392
1	921	120	39	21	62
2	68	3	2	1	0
Subjects with incident AD	996	304	173	101	178
Age at onset, years					
Mean	80.3	84.4	81.1	83.4	81.7
Min	52.5	69	72	65	66
Max	109.5	89	89	90	90
APOE e4 count					
0	471	280	155	87	139
1	424	24	18	14	39
2	101	0	0	0	0

Abbreviations: ACT, Adult Changes in Thought; AD, Alzheimer's disease; ADC, Alzheimer's Disease Center; ADSP, Alzheimer's Disease Sequencing Project; APOE, apolipoprotein E; CHS, Cardiovascular Health Study; RS, Rotterdam Study.

(Fig. 1B–C). Based on these transition ages, we stratified subjects into a younger group and an older group for men and women separately. We estimated that the HRs of e4 allele counts for younger women, younger men, older women, and older men were 6.2, 3.2, 3.8, and 1.8, respectively (Fig. 1D, Table 2). Pairwise differences of the HRs between these groups were statistically significant (z-test *P* values range: $10^{-5} \sim 10^{-3}$). Such stratification also restored the proportionality within each group (goodness-of-fit test *P* value range: 0.22 ~ 0.46).

Next, we examined whether stratification using the transition age of 80.5 years in men and 74.5 years in women would resolve discrepancies among four independent cohorts (Adult Changes in Thought, Cardiovascular Health Study, Framingham Heart Study, and Rotterdam Study). Originally, the HRs of e4 allele counts estimated in these cohorts were contradictory, with some showing harmful effects, some showing protective effects, and others with no statistical significance. After age stratifications, we successfully identified e4 allele as a strong risk factor of AD in the younger men (age < 80.5 years) and younger women (age < 74.5 years) in all cohorts (HR range 4.1 ~ 12.4, Table 2). Except for a borderline *P* value of 0.06 in the Adult Changes in Thought young females, all other *P* values were <0.01. In the older groups, however, the HRs of e4 allele counts were often lower than 1 (HR range: 0.1 ~ 0.5), indicating potential selection biases.

In addition to age stratifications, we explored stratifications on hazards to account for the nonproportionality. We
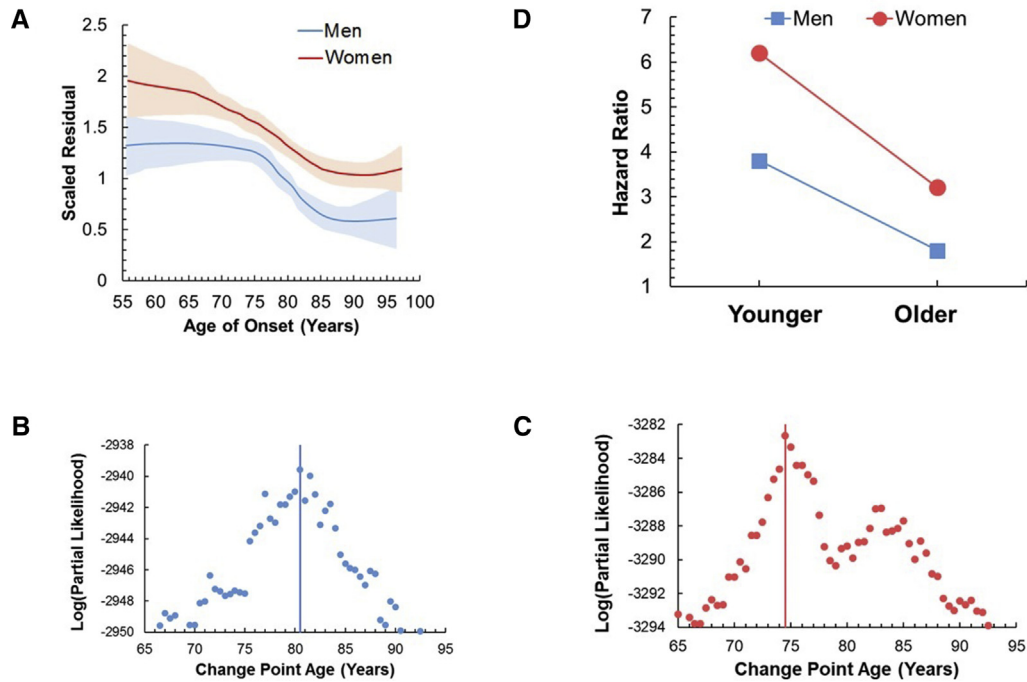
Fig. 1. Age-dependent effect of the APOE e4 allele in the ADC cohort. (A) Smoothed curves of scaled Schoenfeld residuals over age. Shaded areas represent 95% confidence intervals. (B and C) Log partial likelihood of an extended Cox model that uses various age cutoff values to define time-dependent covariates in men (B) and in women (C). (D) Estimated hazard ratios for each sex and age group. In men, the younger group and the older group were defined using cutoff age of 80.5 years. In women, the cutoff age was 74.5 years. Abbreviations: ADC, Alzheimer's Disease Center; APOE, apolipoprotein E.

used a semiparametric approach that treats nonproportional hazards as a result of an LDDP mixture of heterogeneous survival models. The Bayesian analysis found 16 clusters (i.e., models) in men in the ADC data set, among which the top two largest clusters included >95% of the subjects. Similarly, a mixture of six clusters was found in women, among which the top two clusters included >99% of the subjects. Using pointwise posterior estimates of survival probabilities, we computed the corresponding HR of e4 allele counts at each grid point (range: 55 ∼ 65 yo, interval: 1 year). The plots confirmed our previous observations that the HRs declined over time in both men and women and

remained at a lower level after the age of 75–80 years (Fig. 2). These patterns were not sensitive to the priors used in Bayesian analyses.

Interestingly, both the Cox regression and the Bayesian analysis of LDDP mixture models indicated that the HR of e4 allele counts was high at ages younger than 65 years. The ADC data set included subjects diagnosed with AD as early as 52 yo. To assess the effect of the e4 allele in YOAD, we built a Cox regression model to include 758 subjects with baseline age younger than 65 years. Among these subjects, 29 developed AD before the age of 65 years. We estimated that the HRs of e4 allele counts in this

Table 2
Estimated HR and *P* value in Cox regression models for e4 allele

Cohort	All Ages			Younger			Older		
	N	HR	P value	N	HR	P value	N	HR	P value
Male									
ADC	1787 (477)	2.6	0***	1518 (269)	3.8	0***	507 (208)	1.8	10^{-6}***
ACT	500 (114)	1.1	0.7	470 (30)	7.6	10^{-7}***	362 (84)	0.1	0.01**
CHS	258 (58)	2.3	0.004***	227 (31)	4.6	10^{-7}***	167 (27)	0.4	0.3
FHS	154 (37)	2.6	0.02*	143 (11)	12.4	10^{-4}***	101 (26)	0.4	0.4
RS	211 (40)	1.4	0.4	193 (18)	4.1	0.003**	171 (22)	0.2	0.1
Female									
ADC	2940 (519)	3.4	0***	2813 (127)	6.2	0***	1432 (392)	3.2	0***
ACT	634 (190)	0.3	10^{-4}***	630 (4)	4.6	0.06+	440 (186)	0.2	10^{-4}***
CHS	354 (115)	0.4	0.02*	345 (9)	7.4	10^{-4}***	239 (106)	0.0	1.0
FHS	196 (64)	0.8	0.7	191 (5)	5.7	0.01**	131 (59)	0.5	0.2
RS	421 (138)	2.1	10^{-4}***	407 (14)	6.0	10^{-3}***	283 (124)	1.6	0.04*

Abbreviations: ACT, Adult Changes in Thought; CHS, Cardiovascular Health Study; FHS, Framingham Heart Study; RS, Rotterdam Study.
NOTE. Values indicate total number of subjects (number of affected subjects).
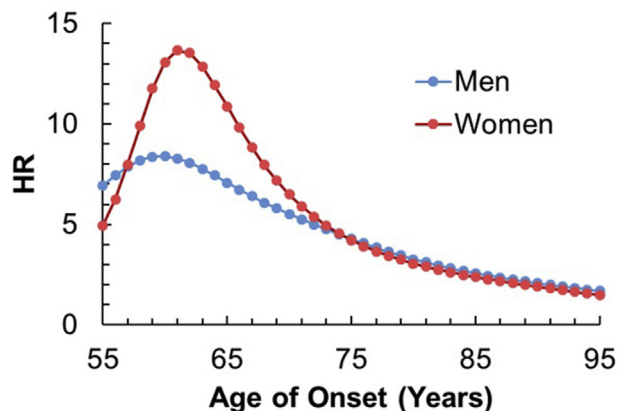NOTE. + *P* value < .1; * *P* value < .05; ** *P* value < .01; *** *P* value < .001.

Fig. 2. Pointwise posterior estimates of HR in the ADC cohort. For each age in a grid (range: 55 ∼ 95 yo, interval: 0.5 years), the corresponding posterior estimate of HR was plotted. Abbreviations: ADC, Alzheimer's Disease Center; HR, hazard ratio.

subcohort were 2.75 for men (*P* value = .003) and 5.3 for women (*P* value = 2×10^{-4}). Therefore, e4 allele was a significant risk factor for YOAD with symptomatic onset before the age of 65 years.

4. Discussion

Since the hallmark discovery associating the APOE genotype with LOAD, it has been commonly assumed that the risk conferred by the e4 allele is constant during aging [4–6]. However, the hypothesis of proportionality has never been tested. In this study, we used two different methods to examine the temporal changes of hazards of e4 allele counts, which consistently revealed a decline with increasing age in both men and women. The assumption of proportionality is seldom true in practice, and we showed that the departure was quite large in all five of the cohorts we examined.

This time-varying pattern can confound analyses in unexpected ways. For example, the HR of e4 allele counts in women is higher than that in men only when ages are comparable (Fig. 1D). If female participants are significantly older than male participants in a study, traditional analysis of the APOE effect may mistakenly report no differences between sexes or a stronger effect in men than in women. This problem can be more severe in underrepresented population subgroups due to small sample sizes [22], likely contributing to the contradictory findings in genetic admixture populations [11,12]. Similarly, in the search for novel genetic factors or biomarkers of AD, unawareness of the age-dependent decline of APOE genotype impact may lead to overestimations or underestimations of the significance and effect size of the target of interest. As large-scale assays enable the screening of targets with small effect sizes and complementary signals, the analyses are more vulnerable to residual biases [23,24]. Thus, our discovery of the age-dependent patterns warrants careful consideration of experimental design and model selection.

In demonstrating that the HRs of e4 allele counts exhibited a stepwise decline in men and women, we propose a sex and age stratification strategy to ensure consistent effects of the APOE genotype within each stratum. Appropriate subject stratification is crucial in experimental design and data analysis. Current age stratification strategies include arbitrary cutoffs ranging from a single ad hoc threshold to 5- or 10-year intervals. Our analysis suggests the optimal cutoffs are closer to 80 yo in men and 75 yo in women. By applying this strategy to four cohorts from the Alzheimer's Disease Sequencing Project study, we identified effects of the APOE genotype that were previously obscured by selection biases.

As to model selection, we examined two semiparametric alternatives of the traditional parametric models. One is an extended Cox regression with time-dependent coefficient, and the other is an LDDP mixture of survival models. Although both approaches can accommodate age-dependent effects effectively and flexibly, they are complementary to each other in applications. In particular, the extended Cox regression allowed us to identify the transition age but introduced a jump in the hazard function and a discontinuity in the estimated HRs. The LDDP analysis enabled adaptive smoothing based on a mixture of heterogeneous models. However, the interpretation of pointwise posterior estimates of HRs was not straightforward. Furthermore, the high computational cost of LDDP analysis may preclude its use in large-scale data analysis. Overviews of common approaches to modeling nonproportional hazards have been published before [25,26].

Finally, we found the effect of the e4 allele was substantial for YOAD whose symptomatic onset is before the age of 65 years. The cause of YOAD in most patients remains unknown. Familial clustering of YOAD cases has led to the discovery of highly penetrant mutations in three genes (*APP*, *PSEN1*, and *PSEN2*) that converge on cerebral amyloid aggregation [27–31], yet collectively, such pathogenic variants explain only a small minority of YOAD cases [32,33]. Contrary to the general view that APOE e4 is a risk factor simply for LOAD, we found it to be a powerful risk factor as well for YOAD, and this finding is consistent with previous studies [8]. As clinical trials migrate to earlier stage disease in younger patients, recruitment strategies will require the identification of such appropriate risk factors as APOE e4.

Limitations of our study should be considered to keep our findings in the proper context. First, the five data sets we analyzed were based on clinical diagnosis, and with only a few exceptions lacked autopsy confirmation. Second, the impact of APOE e4 in older patients, especially those over the age of 80 years could be attenuated by differential survival. In large longitudinal community-based cohorts, APOE e4 has been associated with reduced survival, especially in women, but has not been associated with premature death (before the age of 70 years) [34].

In summary, age compositions of study cohorts can significantly bias the estimated effect of sex and APOE

genotype. If multiple cohorts need to be combined to increase the statistical power, efforts shall be made to harmonize age distributions [35,36]. Furthermore, APOE e4 should be considered a significant risk factor for nonfamilial YOAD and LOAD cases.

Acknowledgments

Authors' contributions: L.L. and R.J.C. conceived the experiment and co-wrote the manuscript. L.L. designed and conducted the analyses.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trci.2018.09.006.

RESEARCH IN CONTEXT

1. Systematic review: We reviewed studies that examined how the interplay of age, sex, and APOE genotype affects the risk of Alzheimer's disease and found inconsistent results. We hypothesized that hazards of the APOE genotype for Alzheimer's disease (AD) are nonproportional, which violates assumptions in commonly used statistical models. Consequently, different age compositions in study cohorts led to biased estimation of disease risk.

2. Interpretation: Our results revealed that APOE genotype–related risks for AD vary significantly with age, and discrepancies in previous reports could be rectified by appropriate age stratification. This strategy also revealed that the APOE e4 allele is a significant risk factor of AD onset before the age of 65 years. Therefore, future experimental design and data analysis should consider estimates of age-dependent risk.

3. Future directions: Our analysis was restricted to subjects with Caucasian ancestry. Given the known health disparities of AD, additional studies should examine whether age-dependent effects are also present in other ethnic groups and if optimal age brackets vary.

References

[1] Qiu C, Kivipelto M, Aguero-Torres H, Winblad B, Fratiglioni L. Risk and protective effects of the APOE gene towards Alzheimer's disease in the Kungsholmen project: variation by age and sex. J Neurol Neurosurg Psychiatry 2004;75:828–33.

[2] Altmann A, Tian L, Henderson VW, Greicius MD, Alzheimer's Disease Neuroimaging Initiative I. Sex modifies the APOE-related risk of developing Alzheimer disease. Ann Neurol 2014;75:563–73.

[3] Riedel BC, Thompson PM, Brinton RD. Age, APOE and sex: Triad of risk of Alzheimer's disease. J Steroid Biochem Mol Biol 2016; 160:134–47.

[4] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993;261:921–3.

[5] Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC Jr, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat Genet 1994;7:180–4.

[6] Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol 2013; 9:106–18.

[7] van Duijn CM, de Knijff P, Cruts M, Wehnert A, Havekes LM, Hofman A, et al. Apolipoprotein E4 allele in a population-based study of early-onset Alzheimer's disease. Nat Genet 1994;7:74–8.

[8] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA 1997; 278:1349–56.

[9] van der Flier WM, Pijnenburg YA, Fox NC, Scheltens P. Early-onset versus late-onset Alzheimer's disease: the case of the missing APOE varepsilon4 allele. Lancet Neurol 2011;10:280–8.

[10] Frisoni GB, Manfredi M, Geroldi C, Binetti G, Zanetti O, Bianchetti A, et al. The prevalence of apoE-epsilon4 in Alzheimer's disease is age dependent. J Neurol Neurosurg Psychiatry 1998; 65:103–6.

[11] Lane KA, Gao S, Hui SL, Murrell JR, Hall KS, Hendrie HC. Apolipoprotein E and mortality in African-Americans and Yoruba. J Alzheimers Dis 2003;5:383–90.

[12] Teruel BM, Rodriguez JJ, McKeigue P, Mesa TT, Fuentes E, Cepero AA, et al. Interactions between genetic admixture, ethnic identity, APOE genotype and dementia prevalence in an admixed Cuban sample; a cross-sectional population survey and nested case-control study. BMC Med Genet 2011;12:43.

[13] Mehta KM, Yeo GW. Systematic review of dementia prevalence and incidence in United States race/ethnic populations. Alzheimer's Demen: J Alzheimer's Assoc 2017;13:72–83.

[14] Neu SC, Pa J, Kukull W, Beekly D, Kuzma A, Gangadharan P, et al. Apolipoprotein E Genotype and Sex Risk Factors for Alzheimer Disease: A Meta-analysis. JAMA Neurol 2017;74:1178–89.

[15] Verghese PB, Castellano JM, Holtzman DM. Apolipoprotein E in Alzheimer's disease and other neurological disorders. Lancet Neurol 2011;10:241–52.

[16] Therneau T, Crowson C, Atkinson E. Using time dependent covariates and time dependent coefficients in the cox model. Red 2013;2:1.

[17] Ungar L, Altmann A, Greicius MD. Apolipoprotein E, gender, and Alzheimer's disease: an overlooked, but potent and promising interaction. Brain Imaging Behav 2014;8:262–73.

[18] Morris JC, Weintraub S, Chui HC, Cummings J, Decarli C, Ferris S, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. Alzheimer Dis Assoc Disord 2006;20:210–6.

[19] Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. Biometrika 1994;81:515–26.

[20] Therneau T. A Package for Survival Analysis in S. R package version 2.37-7; 2014. http://CRAN.R-project.org/package=survival. Accessed September, 2018.

[21] Jara A, Hanson TE, Quintana FA, Muller P, Rosner GL. DPpackage: Bayesian Non- and Semi-parametric Modelling in R. J Stat Softw 2011;40:1–30.

[22] Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14:365–76.

[23] Mahyar UW, Burley JS, Gyllenhaal C, Soejarto DD. Medicinal plants of Seberida (Riau Province, Sumatra, Indonesia). J Ethnopharmacol 1991;31:217–37.

[24] Gosho M, Nagashima K, Sato Y. Study designs and statistical analyses for biomarker research. Sensors (Basel) 2012;12:8966–86.

[25] Hagar Y, Dukic V. Comparison of hazard rate estimation in R; 2015. arXiv preprint arXiv:150903253.

[26] Klein JP, Van Houwelingen HC, Ibrahim JG, Scheike TH. Handbook of survival analysis. UK: Taylor & Francis Group; 2016.

[27] De Strooper B, Saftig P, Craessaerts K, Vanderstichele H, Guhde G, Annaert W, et al. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. Nature 1998;391:387–90.

[28] Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 1991;349:704–6.

[29] Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science 1995;269:973–7.

[30] Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science 1992;258:668–71.

[31] St George-Hyslop PH, Tanzi RE, Polinsky RJ, Haines JL, Nee L, Watkins PC, et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science 1987;235:885–90.

[32] Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, et al. Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. Am J Hum Genet 1999;65:664–70.

[33] Nicolas G, Wallon D, Charbonnier C, Quenez O, Rousseau S, Richard AC, et al. Screening of dementia genes by whole-exome sequencing in early-onset Alzheimer disease: input and lessons. Eur J Hum Genet 2016;24:710–6.

[34] Kulminski AM, Arbeev KG, Culminskaya I, Arbeeva L, Ukraintseva SV, Stallard E, et al. Age, gender, and cancer but not neurodegenerative and cardiovascular diseases strongly modulate systemic effect of the Apolipoprotein E4 allele on lifespan. Plos Genet 2014;10:e1004141.

[35] Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, et al. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. Int J Epidemiol 2010;39:1345–59.

[36] Griffith LE, van den Heuvel E, Raina P, Fortier I, Sohel N, Hofer SM, et al. Comparison of Standardization Methods for the Harmonization of Phenotype Data: An Application to Cognitive Measures. Am J Epidemiol 2016;184:770–8.