

Studying the Interplay Between Apolipoprotein E and Education on Cognitive Decline in Centenarians Using Bayesian Beta Regression

Qingyan Xiang¹, Stacy Andersen², Thomas T Perls², Paola Sebastiani³

¹Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States.

²Section of Geriatrics, Department of Medicine, Boston University School of Medicine, Boston, MA, United States.

³Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, United States.

* Correspondence:

Qingyan Xiang
qyxiang@bu.edu

Keywords: Apolipoprotein E₁, beta regression₂, Blessed Information-Memory-Concentration test₃, cognitive function₄, centenarians₅.

Abstract

Apolipoprotein E (*APOE*) is an important risk factor for cognitive decline and Alzheimer's disease in aging individuals. Among the 3 known alleles of this gene: e2, e3, and e4, the e4 allele is associated with faster cognitive decline and increased risk for Alzheimer's and dementia, while the e2 allele has a positive effect on longevity, and possibly on preservation of cognitive function. Education also has an important effect on cognition and longevity but the interplay between *APOE* and education is not well characterized. Previous studies of the effect of *APOE* on cognitive decline often used linear regression with the normality assumption, which may not be appropriate for analyzing bounded and skewed cognitive test scores. In this paper, we applied Bayesian beta regression to assess the association between *APOE* alleles and cognitive decline in a cohort of centenarians with longitudinal assessment of their cognitive function. The analysis confirmed the negative association between older age and cognition and the beneficial effect of education that persists even at the extreme of human lifespan in carriers of the e3 allele. In addition, the analysis showed an association between *APOE* and cognition that is modified by education. Surprisingly, an antagonistic interaction existed between higher education and *APOE* alleles, suggesting that education may reduce both positive and negative effects of this gene.

1 Introduction

Declines of certain cognitive abilities are common complications of aging and identifying risk factors for cognitive decline is essential to search for therapeutic interventions. Risk factors for cognitive decline include older age, lower education and genes such as apolipoprotein E (*APOE*) that plays an important role in the risk for Alzheimer's disease (Fan et al. 2019). *APOE* is involved in the transport of cholesterol and other lipids between cellular structures (Mahley 1988). The gene has three well characterized alleles e2, e3 and e4 that are defined by the combinations of the genotypes of the single nucleotide polymorphisms rs7412 and rs429358. Studies have shown that e4-carriers have increased risk for Alzheimer's disease and accelerated cognitive decline compared to non-carriers (Stahelin et

al. 1999; Bondi et al. 2003; Wetter et al. 2005), while e2-carriers appears to have a reduced risk for age-related neurodegenerative disease (Henderson et al. 1995; Raber et al. 2004; Kim et al. 2017). The review by O'Donoghue (O'Donoghue et al. 2018) lists 40 studies of the association between *APOE* and cognition in longitudinal studies, but none of these studies examined this association among individuals at extreme ages (e.g., centenarians). Investigating the association of this gene with cognitive decline in centenarians is important for informing about the extent of the protective and deleterious effect of this gene on cognitive function at the extreme of human lifespan. In addition, examining the interaction between education and *APOE* alleles could help to better characterize the long term effect of education on cognition and the interplay between genetic and environmental risk factors. Therefore, in this work, we examined the association of the e2 and e4 alleles of *APOE* with cognitive function in centenarians enrolled in the New England Centenarian Study (NECS) (Sebastiani and Perls 2012), who are enriched for carriers of the e2 allele, have varying levels of education, and for whom we have longitudinally collected assessments of cognitive function.

Typically, studies of the association of *APOE* with cognitive decline use linear mixed models of data collected from a variety of neuropsychological tests (Blair et al. 2005; Caselli et al. 2009; Kim et al. 2017). These models assume unbounded and normally distributed errors, although the outcome of any neuropsychological test is usually defined in a limited interval between 0 and a maximum test value. In addition, the distribution of neuropsychological test scores is often skewed and asymmetric, thus introducing an additional violation of the normality assumption. Violating the normal assumption when modeling cognitive test scores could lead to a poor estimation of the genetic association with cognitive function, and predict scores that are either negative or exceed the maximum test value. To address this problem, we propose using the regression model based on the assumption that the response is beta-distributed. Beta distribution is defined in a finite interval and is flexible to accommodate distribution with different shapes. Hence this distribution is well suited to model bounded test scores and questionnaire outcomes in nature (Smithson and Verkuilen 2006).

Beta regression was proposed by Ferrari and Cribari-Neto (Ferrari and Cribari-Neto 2004) and further developed in the frequentist perspective (Smithson and Verkuilen 2006; Simas et al. 2010) and in the Bayesian perspective (Figueroa-Zúñiga et al. 2013; Cepeda-Cuervo et al. 2016). Studies in multiple fields have utilize beta regression to model different variables such as ischemic stroke lesion volume (Swearingen et al. 2011), genetic distance (Branscum et al. 2007), and understory vegetation communities (Eskelson et al. 2011). In this paper, we describe a Bayesian hierarchical beta regression model to analyze the association between *APOE* and cognitive decline among centenarians enrolled in the NECS (Sebastiani and Perls 2012).

2 Materials and Method

2.1 Participants

The NECS began in 1994 as a population-based study of centenarians living within eight towns in the Boston area and expanded enrollment to include centenarians, their siblings and offspring as well as controls to North America in 2000. Enrolled participants provide socio-demographic data, medical history, and physical function ability (Barthel Index, Mahoney and Barthel 1965; Sinoff and Ore 1997). Participants are administered the Blessed Information-Memory-Concentration (BIMC) test (Blessed et al. 1968; Kawas et al. 1995), a brief test of global cognition that can be administered over the phone to the participant or with the help of a proxy. The BIMC has a maximum total score of 37 points. Scores of 34 or greater represent no impairment, 27 to 33 indicate mild impairment, 21 to 26 signify moderate impairment, and less than 20 are associated with severe impairment (Kawas et

al. 1995). NECS participants are followed annually to collect new medical events, changes in medication and physical function ability, and are administered the BIMC. The NECS study is still open for enrollment. In this analysis, we use the data collected through November 2019. Genotype data for *APOE* were inferred from the combinations of the SNPs s7412 and rs429358 as described in (Sebastiani, Gurinovich, et al. 2019).

2.2 Statistical analysis

2.2.1 Beta regression specification

A variable y defined in the interval $(0, 1)$ follows the beta distribution if the density function is proportional to

$$f(y|\mu, \phi) \propto y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},$$

where μ represents the mean: $0 < \mu < 1$; the function $\mu(1-\mu)/(1+\phi)$ is the variance, and $\phi > 0$ is the precision parameter (Ferrari and Cribari-Neto 2004). To parameterize the mean as a function of covariates, it is convenient to use the logit function

$$g(\mu) = \log(\mu/(1-\mu)) = \mathbf{x}_i^T \boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is a vector of coefficient and \mathbf{x}_i^T is a vector of covariates (Smithson and Verkuilen 2006; Zeileis et al. 2010). The beta distribution is defined in the open interval $(0, 1)$. To fit this model to data defined in the range (a, b) , we use the transformation $y' = (y - a)/(b - a)$. In our analysis, the BIMC scores range from 0 to 37 and we rescaled the data to the interval $[0.01, 0.99]$ to avoid zeros and ones.

2.2.2 Bayesian beta regression modelling

We were interested in modelling the association of the following main covariates with BIMC scores: age, sex, education and *APOE* alleles. We standardized the continuous variables *age* and *education* to generate parameters on the same scale and coded the dichotomous variable *sex* 0 for female and 1 for male. Since homozygote carriers of e2 or e4 are rare, we used the following allele groupings in the analyses:

- “e2 group” comprising carriers of the *APOE* genotypes e2e2 or e2e3;
- “e3 group” comprising carriers of the *APOE* genotype e3e3;
- “e4 group” comprising carriers of the *APOE* genotypes e3e4 or e4e4.

And we selected “e3 group” as the reference group because it is the most frequent genotype in Whites. We performed two analyses to distinguish the beneficial association of the e2 allele relative to e3 from lack of carrying the deleterious e4 allele. One analysis estimated the association between *APOE* and BIMC scores of the e2 group relative to the e3 group, including only the e2 and e3 groups. The other analysis estimated the association between *APOE* and BIMC scores of the e4 group relative to e3 group.

We used backward selection for model fitting, and started with the hierarchical regression model:

$$y_{ij} \sim \text{Beta}(\mu_{ij}\phi, \phi(1-\mu_{ij}))$$

$$\text{logit}(\mu_{ij}) = \mu_{b_0} \cdot I_{single} + b_{i0} \cdot (1 - I_{single}) + \beta_{age} \cdot age_{ij} + \beta_{sex} \cdot sex_i + \beta_{edu} \cdot edu_i + \beta_{APOE} \cdot APOE_i + \beta_{age.edu} \cdot age_i \cdot edu_i + \beta_{age.APOE} \cdot age_i \cdot APOE_i + \beta_{edu.APOE} \cdot edu_i \cdot APOE_i,$$

where y_{ij} denotes the j^{th} cognitive test score of the i^{th} participant, the β coefficients are fixed effects and b_{i0} is the random intercept that we used to account for within participant correlation of the repeated measurements. We used a piecewise random intercept $\mu_{b_0} \cdot I_{single} + b_{i0} \cdot (1 - I_{single})$ to accommodate for participants with different number of test administration, where the indicator variable I_{single} is 0 for participants with only one cognitive test score and $I_{single} = 1$ otherwise. The random intercepts b_{i0} were assumed to be independent and normally distributed, i.e., $b_{i0} \sim N(\mu_b, \sigma_b^2)$. We specified a normal prior for the mean parameter μ_b of the random intercept that $\mu_b \sim N(0, 1000)$, and a gamma distribution with the shape and scale parameters both equal to 1 for the precision parameter ϕ and the precision of random intercept $1/\sigma_b^2$. With this parameterization, the participants with only one test score would be assigned the fixed intercept μ_b in the regression, while participants with more than one test score have their own random intercept b_{i0} . We used normal priors for all the fixed effect parameters $\beta \sim N(0, 1000)$.

We implemented the backward model selection algorithm using the deviance information criterion (DIC) (Spiegelhalter et al. 2002; 2014), which is particularly useful for selection of hierarchical models. Since DIC has a tendency to overfit (Clarke and Clarke 2018), we then refined the model selected by this search by retaining only interactions and main effects with a posterior credible interval that did not include 0. Once we selected the final model, we also conducted a sensitivity analysis with respect to the prior distributions. For regression coefficients and the mean parameter of the random intercept that use the normal priors, we reduced the variance from 1000 to 10 and 100. For precision parameter of the beta distribution and the reciprocal of the variance of the random intercept that use the gamma priors, we modified the variance of gamma priors from 1 to 100. We ran each case to assess if the parameter estimates altered after we changed the scale of the prior parameters. All analyses were conducted in R3.6 and all Bayesian models were analyzed using Markov Chain Monte Carlo (MCMC) implemented in the “rjags” package (Plummer 2016). The posterior estimates of the parameters derived from at least 8000 burn-in adaptations and 4000 iterations.

2.2.3 Interpretation of the results

A limitation of beta regression is that the magnitude of the regression coefficients is not directly interpretable in terms of changes of the outcome. To better understand the association of *APOE* with BIMC scores and the interplay with education, we estimated the fitted means using inverse transformation of the logit function, and then we rescaled the fitted means to the original scale of the BIMC scores (0, 37). We calculated fitted means for the e2, e3 and e4 groups with respect to three different education levels: low education (25% quantile of study population’s education; 8 years), median education (median of study population’s education; 12 years) and high education (75% quantile of study population’s education; 15 years). For each allele in each education level, we also calculated the corresponding age of onset of moderate cognitive impairment (BIMC score = 26) using the predicted trajectories of the cognitive test scores. These ages provide a quantitatively more interpretable metric of the genetic effects.

3 Results

Out of 768 total participants in the NECS dataset with *APOE* genotypes, we excluded 167 participants with missing test scores, 111 participants with missing education information, and 4 participants with *APOE* genotype e2e4. Table 1 summarizes the demographic characteristics and test scores at baseline of the remaining 486 participants used in our analysis. The *APOE* e3 group was the most prevalent and used as the reference group in all subsequent analyses. Approximately 24% of centenarians in the study carried at least one e2 allele, while only 8% were carriers of an e4 allele. We did not find any e4 homozygous centenarians.

The histogram in Figure 1 shows that the distribution of the baseline BIMC scores is bounded and highly skewed. Hence we analyzed the rescaled test scores using beta regression. The combination of backward selection using the DIC criterion and refinement of the model using the posterior credible intervals produced the models summarized in Tables 2 and 3. The diagnostic plots (trace plots, autocorrelation plots and the Gelman plots of all MCMC chains) showed no indications of lack of convergence. The sensitivity analyses also showed that changing the prior distributions did not appreciably alter the parameter estimates. The diagnostic plots and sensitivity analyses are all included in the supplementary files.

Table 2 summarizes the results of the analysis comparing *APOE* e2 with *APOE* e3, while Table 3 summarizes the results of the analysis comparing *APOE* e4 with *APOE* e3. The significant coefficients are marked with bold fonts. In both analyses, age, sex and education were all significantly associated with BIMC scores. Older age was associated with significantly worse performance on the test. More years of education and male sex were associated with significantly higher scores.

In the analysis comparing e2 with e3, the interaction between education and e2 was significant, thus suggesting an effect modification of education on the association of e2 with BIMC scores. It is noteworthy that the estimates of the main effect of education (0.063) and the interaction term (-0.062) were opposite, so that the negative interaction between e2 and education numerically canceled out the positive association of education with BIMC scores in carriers of the e2 allele. This is shown in Figure 2, in which the fitted BIMC scores in e2 carriers remain nearly the same in different education levels. Figure 2 also shows the positive association of education with BIMC score in the e3 group. Interestingly, only participants of the e3 group with low education had a worse BIMC score compared to carriers of one or more e2 alleles. The difference between the e2 and e3 groups became negligible in participants with median education level, and carriers of the e3e3 genotypes with a high level of education obtained a better score than carriers of the e2 alleles.

To better summarize the clinical implication of these results, we calculated the ages of onset of cognitive impairment predicted by the fitted model. From the fitted lines, we estimated that the ages of onset of moderate cognitive impairment were 99.5, 101.8 and 103.5 years in for e3 carriers with low, median and high education groups, respectively. The age of onset of moderate cognitive impairment (BIMC score = 26) was 102.1 years in e2 carriers, independent of education. Therefore, e2 carriers were estimated to delay the onset of moderate cognitive impairment by approximate 2 years compared to e3 carriers with low education, but this advantage essentially disappears with higher education.

In the analysis comparing e4 with e3, both the main effect term of e4 and the interaction with education were significantly negative, thus suggesting that higher education was not sufficient to remove the negative association of the e4 allele with BIMC scores. This is illustrated in Figure 3 that shows the predicted BIMC scores in the e3 and e4 groups stratified by education. Among individuals

with the e3e3 genotype, more years of education was associated with higher BIMC scores, but the positive effect modification of higher education was reduced in carriers of e4 compared to e3.

From the fitted lines, the ages of onset of moderate cognitive impairment in e4 carriers were 99.5, 98.9 and 98.57 years in participants with low, median and high education respectively. Hence, in centenarians carrying the e4 allele, more years of education was not associated with a delay of cognitive decline.

4 Discussion

In this paper, we investigated the relationship between *APOE* alleles and change of cognitive function in a large cohort of centenarians enriched for carriers of the e2 allele. We used a Bayesian hierarchical beta regression model to characterize the association of *APOE* with cognitive function assessed through the Blessed Information-Memory-Concentration test. This method allowed us to fit the test scores in the range of admissible values, and to model non-linear relations of the score with age and education. Our analyses confirmed the decline of cognitive function and the positive association of education with preservation of cognitive function at extreme old age. The analysis also showed that the allele e4 has a negative association with cognitive function even at extreme old ages, while the allele e2 appears to be protective only in centenarians with low education. On the other hand, higher education appeared to increase the magnitude of the association between the e4 alleles and poorer cognitive function.

Our findings are consistent with previous studies that showed a decline of cognitive function with older age in multiple domains (Harada et al. 2013), and the importance of early education on better cognitive function at older age (Alley et al. 2007). Longitudinal studies of cognitive function in a large cohort of centenarians are limited. A study of Japanese centenarians detected an education by *APOE* e4 interaction on cognition that differed by sex (Ishioka et al. 2016). The study showed that highly educated centenarians who carried at least one e4 allele had worse performance in the Mini-Mental State Examination, compared to poorly educated centenarians. Our results are consistent with this counter-intuitive result but go one step further and disentangle the negative association of the e4 allele from the positive association of the e2 allele.

There is some evidence that the e2 allele confers protection from cognitive decline (Kim et al. 2017), in addition to increasing the chance for longevity (Sebastiani, Gurinovich, et al. 2019) and conferring protection from aging-related diseases (Wolters et al. 2019; Kuo et al. 2020). For example, we identified a cluster of participants in the Long Life Family study who have slower decline of processing speed and attention and are genetically enriched for the e2 allele (Sebastiani et al. 2020). These results suggest that targeting e2 gene products could lead to important therapeutics (Sebastiani, Monti, et al. 2019). The interaction between education and the e2 allele is also consistent with the fact that more years of education is associated with better cognitive function in later life (Alley et al., 2007; Wilson et al., 2009). Given the heterogeneity of the association between the e2 allele and longevity in populations with different genetic backgrounds (Gurinovich et al. 2019), it would be interesting to extend these analyses to also include the effect of nutrition on preservation of cognitive function as a possible effect modifier of *APOE* alleles.

The method we used in this analysis has several strengths. First, the cognitive test scores are bounded and skewed. Instead of applying an inadequate regression model assuming normal errors, we used Bayesian beta regression as a more appropriate modeling framework. Second, to increase statistical power, we included all participants with at least one cognitive function assessment in our model. A

piecewise random intercept was used to include subjects with either repeated measurements or just one measurement. Finally, we carried out a model selection using the DIC criterion and then refined our model by credible intervals (CI). From the diagnostic plots of the MCMC, the final models converged well. Finally, in previous studies (Kulminski et al. 2015; Barral et al. 2017) *APOE* genotype was characterized as carriers of e4 (i.e. e2e4, e3e4, and e4e4), and non-carriers of e4 (i.e. e2e2, e2e3, and e3e3). This comparison could yield biased results since the non-carriers of e4 group might not represent the most prevalent population. To eliminate this potential bias, we divided the participants into subgroups, where we only compared the e2 carriers to e3 carriers, and the e4 carriers to e3 carriers.

Notable limitations of this study include loss to follow-up and mortality. Due to the extreme age of the study population, the mortality rate is considerably high: 49% of participants included in our analysis only completed the baseline evaluation. The absence of longitudinal data may induce bias, and thus we used a piecewise random intercept to compensate for those only with one record.

In conclusion, we examined the association between *APOE* alleles and cognitive decline in a cohort of centenarians. We confirmed the negative correlation between the e4 allele and cognition even at the extreme of human lifespan, and newly found the carrying the e2 alleles appears to be beneficial only in centenarians with poor education. The interaction between *APOE* e4 and education produced a counter-intuitive result that is however consistent with other results in centenarians. The antagonistic relation between higher education and carrying the e4 allele may be confounded by other risk factor and warrants more in depth studies.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

Author Q. X. proposed the method used in the paper and conduct the statistical analysis and drafted the manuscript.

Author S. A. and T. P. designed the study and collect data from study cohort and helped editing the manuscript.

Author P. S. designed the analytic strategy, supervised the work and critically revised the manuscript.

7 Funding

This work was supported by the National Institute on Aging (NIA cooperative agreement U19-AG023122 and R01 AG061844).

8 Reference

- Alley, D., Suthers, K., and Crimmins, E. 2007. "Education and Cognitive Decline in Older Americans." *Research on Aging* 29 (1): 73-94. <https://doi.org/10.1177/0164027506294245>.
<https://dx.doi.org/10.1177/0164027506294245>.
- Barral, S., Singh, J., Fagan, E., Cosentino, S., Andersen-Toomey, S. L., Wojczynski, M. K., et al. 2017. "Age-Related Biomarkers in LLFS Families With Exceptional Cognitive Abilities."

The Journals of Gerontology: Series A 72 (12): 1683-1688.

<https://doi.org/10.1093/gerona/glx034>. <https://dx.doi.org/10.1093/gerona/glx034>.

- Blair, C., Folsom, A., Knopman, D. S., Bray, M., Mosley, T., and Boerwinkle, E. 2005. "APOE genotype and cognitive decline in a middle-aged cohort." *Neurology* 64 (2): 268-276.
- Blessed, G., Tomlinson, B. E., and Roth, M. 1968. "The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects." *The British journal of psychiatry* 114 (512): 797-811.
- Bondi, M. W., Houston, W. S., Salmon, D. P., Corey-Bloom, J., Katzman, R., Thal, L. J., et al. 2003. "Neuropsychological deficits associated with Alzheimer's disease in the very-old: Discrepancies in raw vs. standardized scores." *Journal of the International Neuropsychological Society* 9 (5): 783-795. <https://doi.org/10.1017/s1355617703950119>. <https://dx.doi.org/10.1017/s1355617703950119>.
- Branscum, A. J., Johnson, W. O., and Thurmond, M. C. 2007. "Bayesian beta regression: applications to household expenditure data and genetic distance between foot-and-mouth disease viruses." *Australian & New Zealand Journal of Statistics* 49 (3): 287-301.
- Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., Ahern, G. L., et al. 2009. "Longitudinal Modeling of Age-Related Memory Decline and the APOE ϵ 4 Effect." *New England Journal of Medicine* 361 (3): 255-263. <https://doi.org/10.1056/nejmoa0809437>. <https://dx.doi.org/10.1056/NEJMoa0809437>.
- Cepeda-Cuervo, E., Jaimes, D., Marín, M., and Rojas, J. 2016. "Bayesian beta regression with Bayesianbetareg R-package." *Computational Statistics* 31 (1): 165-187.
- Clarke, B. S., and Clarke, J. L. 2018. *Predictive Statistics: Analysis and Inference beyond Models*. Vol. 46. Cambridge University Press.
- Eskelson, B. N., Madsen, L., Hagar, J. C., and Temesgen, H. 2011. "Estimating riparian understory vegetation cover with beta regression and copula models." *Forest Science* 57 (3): 212-221.
- Fan, J., Tao, W., Li, X., Li, H., Zhang, J., Wei, D., et al. 2019. "The Contribution of Genetic Factors to Cognitive Impairment and Dementia: Apolipoprotein E Gene, Gene Interactions, and Polygenic Risk." *International Journal of Molecular Sciences* 20 (5): 1177. <https://doi.org/10.3390/ijms20051177>. <https://dx.doi.org/10.3390/ijms20051177>.
- Ferrari, S., and Cribari-Neto, F. 2004. "Beta regression for modelling rates and proportions." *Journal of applied statistics* 31 (7): 799-815.
- Figueroa-Zúñiga, J. I., Arellano-Valle, R. B., and Ferrari, S. L. P. 2013. "Mixed beta regression: A Bayesian perspective." *Computational Statistics & Data Analysis* 61: 137-147. <https://doi.org/10.1016/j.csda.2012.12.002>. <https://dx.doi.org/10.1016/j.csda.2012.12.002>.
- Gurinovich, A., Andersen, S. L., Puca, A., Atzmon, G., Barzilai, N., and Sebastiani, P. 2019. "Varying Effects of APOE Alleles on Extreme Longevity in European Ethnicities." *The Journals of Gerontology: Series A* 74 (Supplement_1): S45-S51. <https://doi.org/10.1093/gerona/glz179>. <https://dx.doi.org/10.1093/gerona/glz179>.
- Harada, C. N., Natelson Love, M. C., and Triebel, K. L. 2013. "Normal Cognitive Aging." *Clinics in Geriatric Medicine* 29 (4): 737-752. <https://doi.org/10.1016/j.cger.2013.07.002>. <https://dx.doi.org/10.1016/j.cger.2013.07.002>.
- Henderson, A. S., Jorm, A. F., Korten, A. E., Christensen, H., Jacomb, P. A., Easteal, S., et al. 1995. "Apolipoprotein E allele ϵ 4, dementia, and cognitive decline in a population sample." *The Lancet* 346 (8987): 1387-1390. [https://doi.org/10.1016/s0140-6736\(95\)92405-1](https://doi.org/10.1016/s0140-6736(95)92405-1). [https://dx.doi.org/10.1016/S0140-6736\(95\)92405-1](https://dx.doi.org/10.1016/S0140-6736(95)92405-1).
- Ishioka, Y. L., Gondo, Y., Fuku, N., Inagaki, H., Masui, Y., Takayama, M., et al. 2016. "Effects of the APOE ϵ 4 allele and education on cognitive function in Japanese centenarians." *AGE* 38 (5-6): 495-503. <https://doi.org/10.1007/s11357-016-9944-8>. <https://dx.doi.org/10.1007/s11357-016-9944-8>.

- Kawas, C., Karagiozis, H., Resau, L., Corrada, M., and Brookmeyer, R. 1995. "Reliability of the blessed telephone information-memory-concentration test." *Journal of geriatric psychiatry and neurology* 8 (4): 238-242.
- Kim, Y. J., Seo, S. W., Park, S. B., Yang, J. J., Lee, J. S., Lee, J., et al. 2017. "Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: A three-year longitudinal study." *Scientific Reports* 7 (1). <https://doi.org/10.1038/s41598-017-02046-y>. <https://dx.doi.org/10.1038/s41598-017-02046-y>.
- Kulminski, A. M., Arbeev, K. G., Culminskaya, I., Ukraintseva, S. V., Stallard, E., Province, M. A., et al. 2015. "Trade-Offs in the Effects of the Apolipoprotein E Polymorphism on Risks of Diseases of the Heart, Cancer, and Neurodegenerative Disorders: Insights on Mechanisms from the Long Life Family Study." *Rejuvenation Research* 18 (2): 128-135. <https://doi.org/10.1089/rej.2014.1616>. <https://dx.doi.org/10.1089/rej.2014.1616>.
- Kuo, C.-L., Pilling, L. C., Atkins, J. L., Kuchel, G. A., and Melzer, D. 2020. "ApoE e2 and aging-related outcomes in 379,000 UK Biobank participants." *Aging* 12 (12): 12222-12233. <https://doi.org/10.18632/aging.103405>. <https://dx.doi.org/10.18632/aging.103405>.
- Mahley, R. 1988. "Apolipoprotein E: cholesterol transport protein with expanding role in cell biology." *Science* 240 (4852): 622-630. <https://doi.org/10.1126/science.3283935>. <https://dx.doi.org/10.1126/science.3283935>.
- Mahoney, F. I., and Barthel, D. W. 1965. "Functional evaluation: the Barthel Index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill." *Maryland state medical journal*.
- O'Donoghue, M. C., Murphy, S. E., Zamboni, G., Nobre, A. C., and Mackay, C. E. 2018. "APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: A review." *Cortex* 104: 103-123. <https://doi.org/10.1016/j.cortex.2018.03.025>. <https://www.ncbi.nlm.nih.gov/pubmed/29800787>.
- Plummer, M. 2016. "rjags: Bayesian graphical models using MCMC." *R package version* 4 (6).
- Raber, J., Huang, Y., and Ashford, J. W. 2004. "ApoE genotype accounts for the vast majority of AD risk and AD pathology." *Neurobiol Aging* 25 (5): 641-50. <https://doi.org/10.1016/j.neurobiolaging.2003.12.023>. <https://www.ncbi.nlm.nih.gov/pubmed/15172743>.
- Sebastiani, P., Andersen, S. L., Sweigart, B., Du, M., Cosentino, S., Thyagarajan, B., et al. 2020. "Patterns of multi-domain cognitive aging in participants of the Long Life Family Study." *GeroScience*. <https://doi.org/10.1007/s11357-020-00202-3>. <https://dx.doi.org/10.1007/s11357-020-00202-3>.
- Sebastiani, P., Gurinovich, A., Nygaard, M., Sasaki, T., Sweigart, B., Bae, H., et al. 2019. "APOE Alleles and Extreme Human Longevity." *The Journals of Gerontology: Series A* 74 (1): 44-51. <https://doi.org/10.1093/gerona/gy174>. <https://dx.doi.org/10.1093/gerona/gy174>.
- Sebastiani, P., Monti, S., Morris, M., Gurinovich, A., Toshiko, T., Andersen, S. L., et al. 2019. "A serum protein signature of APOE genotypes in centenarians." *Aging Cell* 18 (6): e13023. <https://doi.org/10.1111/accel.13023>. <https://dx.doi.org/10.1111/accel.13023>.
- Sebastiani, P., and Perls, T. T. 2012. "The genetics of extreme longevity: lessons from the New England Centenarian Study." *Front Genet* 3: 277. <https://doi.org/10.3389/fgene.2012.00277>. <http://www.ncbi.nlm.nih.gov/pubmed/23226160>.
- Simas, A. B., Barreto-Souza, W., and Rocha, A. V. 2010. "Improved estimators for a general class of beta regression models." *Computational Statistics & Data Analysis* 54 (2): 348-366. <https://doi.org/10.1016/j.csda.2009.08.017>. <https://dx.doi.org/10.1016/j.csda.2009.08.017>.
- Sinoff, G., and Ore, L. 1997. "The Barthel Activities of Daily Living Index: Self-Reporting Versus Actual Performance in the Old-Old (≥ 75 years)." *Journal of the American Geriatrics Society*

- 45 (7): 832-836. <https://doi.org/10.1111/j.1532-5415.1997.tb01510.x>.
<https://dx.doi.org/10.1111/j.1532-5415.1997.tb01510.x>.
- Smithson, M., and Verkuilen, J. 2006. "A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables." *Psychological Methods* 11 (1): 54-71.
<https://doi.org/10.1037/1082-989x.11.1.54>. <https://dx.doi.org/10.1037/1082-989x.11.1.54>.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. 2002. "Bayesian measures of model complexity and fit." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (4): 583-639. <https://doi.org/10.1111/1467-9868.00353>.
<https://dx.doi.org/10.1111/1467-9868.00353>.
- . 2014. "The deviance information criterion: 12 years on." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76 (3): 485-493. <https://doi.org/10.1111/rssb.12062>.
<https://dx.doi.org/10.1111/rssb.12062>.
- Stahelin, H. B., Perrig-Chiello, P., Mittrache, C., Miserez, A. R., and Perrig, W. J. 1999. "Apolipoprotein E genotypes and cognitive functions in healthy elderly persons." *Acta Neurologica Scandinavica* 100 (1): 53-60. <https://doi.org/10.1111/j.1600-0404.1999.tb00723.x>.
<https://dx.doi.org/10.1111/j.1600-0404.1999.tb00723.x>.
- Swearingen, C. J., Tilley, B. C., Adams, R. J., Rumboldt, Z., Nicholas, J. S., Bandyopadhyay, D., et al. 2011. "Application of Beta Regression to Analyze Ischemic Stroke Volume in NINDS rt-PA Clinical Trials." *Neuroepidemiology* 37 (2): 73-82. <https://doi.org/10.1159/000330375>.
<https://dx.doi.org/10.1159/000330375>.
- Wetter, S. R., Delis, D. C., Houston, W. S., Jacobson, M. W., Lansing, A., Cobell, K., et al. 2005. "Deficits in inhibition and flexibility are associated with the APOE-E4 allele in nondemented older adults." *J Clin Exp Neuropsychol* 27 (8): 943-52.
<https://doi.org/10.1080/13803390490919001>.
<https://www.ncbi.nlm.nih.gov/pubmed/16207619>.
- Wolters, F. J., Yang, Q., Biggs, M. L., Jakobsdottir, J., Li, S., Evans, D. S., et al. 2019. "The impact of APOE genotype on survival: Results of 38,537 participants from six population-based cohorts (E2-CHARGE)." *PLOS ONE* 14 (7): e0219668.
<https://doi.org/10.1371/journal.pone.0219668>.
<https://dx.doi.org/10.1371/journal.pone.0219668>.
- Zeileis, A., Cribari-Neto, F., Grün, B., and Kosmidis, I. 2010. "Beta regression in R." *Journal of statistical software* 34 (2): 1-24.

9 Tables

Table 1. Demographic characteristics and the BIMC scores at baseline of the participants in the New England Centenarian Study

	<i>APOE</i> e2 (e2e2,e2e3)	<i>APOE</i> e3 (e3e3)	<i>APOE</i> e4 (e3e4)
N	117	331	38
Age at Enrollment, mean (SD)	103.54 ± 4.59	103.32 ± 4.43	102.42 ± 4.97
Gender, male (%)	22 (18.8%)	85 (25.68%)	9 (23.68%)
Years of Education, mean (SD)	11.63 ± 3.51	11.82 ± 3.85	12.29 ± 4.05
BIMC scores at baseline, mean (SD)	25.11 ± 8.23	25.00 ± 8.63	24.38 ± 10.20

Table 2. Parameter estimates and 95% credible intervals from the analysis of the BIMC scores in carriers of *APOE* e2 and e3 Alleles

	Estimates	2.5%CI	97.5%CI	SD
Intercept (μ_b)	0.592	0.489	0.695	0.053
Age	-0.110	-0.138	-0.099	0.009
Sex, male	0.262	0.076	0.447	0.096
Education	0.062	0.039	0.085	0.012
<i>APOE2</i>	0.037	-0.162	0.186	0.088
<i>APOE2</i> *edu	-0.063	-0.109	-0.019	0.023

All covariates were standardized and the score was rescaled to the interval (0, 1) to fit the beta regression, so all parameters should be interpreted on the logit scale and the effects are relative to the rescaled score.

Table 3. Parameter estimates and 95% credible intervals from the analysis of the BIMC scores in carriers of *APOE* e4 and e3 Alleles

	Estimates	2.5%CI	97.5%CI	SD
Intercept (μ_b)	0.647	0.541	0.685	0.053
Age	-0.119	-0.141	-0.098	0.011
Sex, male	0.215	0.019	0.413	0.101
Education	0.065	0.041	0.089	0.012
<i>APOE4</i>	-0.382	-0.664	-0.103	0.139
<i>APOE4</i> *edu	-0.082	-0.159	-0.008	0.034

All covariates were standardized and the score was rescaled to the interval (0, 1) to fit the beta regression, so all parameters should be interpreted on the logit scale and the effects are relative to the rescaled score.

10 Figures

Figure 1. Histogram of baseline Blessed Information-Memory-Concentration scores from the New England Centenarian Study.

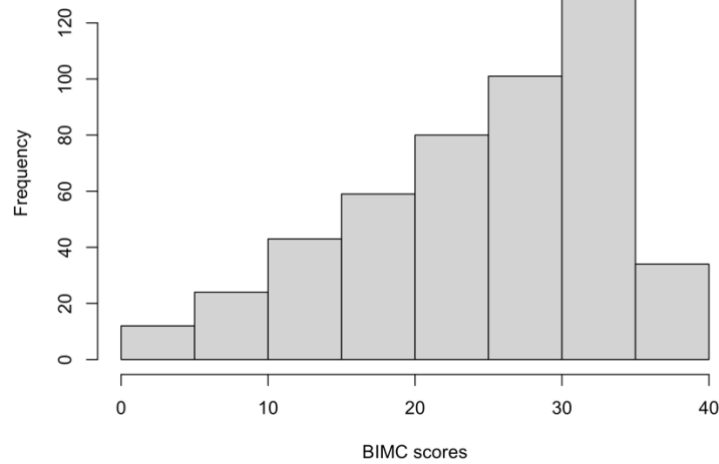


Figure 2. Predicted Blessed Information-Memory-Concentration scores in carriers of one or more e2 alleles (dashed lines) compared to homozygotes e3e3 (continuous lines) for low education (25% quantile, 8 years, red), median education level (12 years, black), and high education level (75% quantile, 15 years, blue). The scores were predicted using the estimates from Bayesian beta regression, assuming sex = female. Note visually there is only one dashed line, because the fitted lines of e2 carriers overlap in all three education groups.

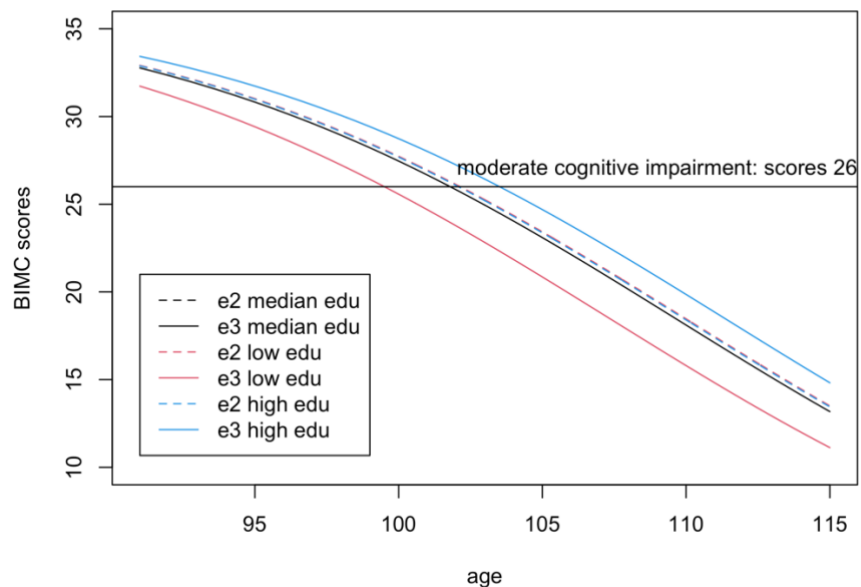


Figure 3. Predicted Blessed Information-Memory-Concentration test scores in carriers of one or more e4 alleles (dashed lines) compared to homozygotes e3e3 (continuous lines) for low education (25% quantile, 8 years, red), median education level (12 years, black), and high education level (75% quantile, 15 years, blue). The scores were predicted using the estimates from Bayesian beta regression, assuming sex = female.

