**Methods**

Whole genome sequencing of 576 individuals over 111 families from Alzheimer’s disease sequencing project (ADSP) was used in the GWAS. In ADSP, each individual’s AD status was clinically examined in one of four levels: no (80), possible (81), probable (360), and definite (55). Additional sample information includes sex, age, and APOE genotypes on three APOE allele types: e2/e3/e4. Frequencies of APOE e2, e3, and e4 are 4%, 78%, and 18%, respectively. Homozygous APOE e2 and APOE e4 are rare (2 and 0 samples). Interquartile range of sample ages is 67 to 80. This age range is typical for late-onset AD; hereby constrain our study in LOAD. Interestingly, while 61% of the samples are female, female proportions in the definite and probable AD groups are higher, 65.5% and 65.6%, respectively (Figure 1).

To detect significant variants for late-onset Alzheimer’s disease (LOAD), we built a generalized linear mixed model (GLMM). AD levels were modeled by an ordered categorical variable. Probability for an individual to fall in the th (j = 1, 2, 3) and any lower categories follows: , where denotes the probability that the th individual falls in category . was logit-transformed and denoted as: , where provides each AD level a unique intercept. is a vector of explanatory variables. is the corresponding effect sizes vector. is a multivariate variable that follows: , with covariance matrix the genotype-based relatedness between individual pairs (IBS). , where is variant number, and is the genotype of individual *i* and *j* on variant *m.* We used u as a random term to account for population structure in the GWAS. 23 kinship matrices were computed by the taking-one-off strategy, in that for any given variant of one chromosome, the corresponding kinship matrix was computed by taking off all variants of the given chromosome. KING was used for fast kinship estimation on the massive genotype data (Manichaikul et al., 2010).

The GLMM was built in Bayesian framework and implemented with Stan (<http://mc-stan.org>). A non-informative prior distribution, *cauchy(0, 1),* was assigned for each parameter. Point estimations of model parameters were obtained by maximizing model’s joint posterior likelihood. Full posterior distributions of model parameters were obtained by sampling using the No-U-Turn sampler (Hoffman and Gelman 2011).

To identify genomic variants that explain the most variance, log-likelihood ratio test (LRT) was computed by subtracting the posterior log-likelihood of the null model from that of the full model for each variant. A total of 22 null models were estimated using the same covariates (age, sex, APOE/e2, APOE/e4) together with one of the 22 taking-one-off kinship matrices. LOD score was approximated as two times the LRT. Significant LOD value was determined empirically by permutation as follow: (1) randomly choose 1 million SNP; (2) randomly permute the chosen SNP over samples; (3) estimate the 1 million models with each permutated SNP and compute the LRT; (4) save the maximal LRT value; (5) repeat steps 1-4 3000 times. This leads to 3000 maximal LRT values. Significant threshold was set as 0.95 quantile of the 3000 maximal LRT values.

A full Bayesian sampling was applied on 193,676 variants that fall in 500 kilo-bp ranges of the 73 significant peaks by maximal likelihood estimation. WAIC (Watanable-Akaike information criterion) was quantified as -2 \* elpd (expected log pointwise predictive density) for each variant to describe model fitness after Bayesian sampling, where (Vehtari and Gelman, 2014).

Three features can characterize our model: (1) the response was modeled by an ordered categorical variable, (2) population structure was modeled by a random multivariate variable with known covariance and estimated by each variant independently, (3) Bayesian framework and inference was applied. To validate the model results, we modified our model to fit established inference tools to compare the results. We constructed a linear mixed model (LMM): , where *N(AD)* was numerical by transforming the categorical AD status: 0/no, 0.25/possible, 0.5/probable, and 1/definite, , . The LMM model was estimated with QTLRel in R (Cheng et al., 2011).

R was used for most analysis except specified (www.r-project.org). Genomic variants were processed with Unix shell, Python, and Plink (Purcell et al., 2007). Variant effects were estimated with Ensembl variant effect predictor (VEP).