Published online: 27 September 2021

DOI: 10.1002/trc2.12211

Alzheimer's & Dementia
Translational Research & Clinical Interventions

RESEARCH ARTICLE

GenoRisk: A polygenic risk score for Alzheimer's disease

Samuel P. Dickson[1] | **Suzanne B. Hendrix**[1] | **Bruce L. Brown**[2] | **Perry G. Ridge**[4] | **Jessie Nicodemus-Johnson**[1] | **Marci L. Hardy**[3] | **Allison M. McKeany**[3] | **Steven B. Booth**[3] | **Ryan R. Fortna**[3] | **John S.K. Kauwe**[2,4] | **NIAGADS**[1]

[1] Pentara Corporation, Salt Lake City, Utah, USA

[2] Department of Psychology, Brigham Young University, Provo, Utah, USA

[3] ADx Healthcare, Bellingham, Washington, USA

[4] Department of Biology, Brigham Young University–Hawaii, Laie, Hawaii, USA

Correspondence
Samuel P. Dickson, Pentara Corporation, 2261 E. 3300 S., Salt Lake City, UT 84109, USA.
E-mail: sdickson@pentaracorporation.com

Abstract

Introduction: Recent clinical trials are considering inclusion of more than just apolipoprotein E (*APOE*) ε4 genotype as a way of reducing variability in analysis of outcomes.

Methods: Case-control data were used to compare the capacity of age, sex, and 58 Alzheimer's disease (AD)–associated single nucleotide polymorphisms (SNPs) to predict AD status using several statistical models. Model performance was assessed with Brier scores and tenfold cross-validation. Genotype and sex × age estimates from the best performing model were combined with age and intercept estimates from the general population to develop a personalized genetic risk score, termed age, and sex-adjusted GenoRisk.

Results: The elastic net model that included age, age x sex interaction, allelic *APOE* terms, and 29 additional SNPs performed the best. This model explained an additional 19% of the heritable risk compared to *APOE* genotype alone and achieved an area under the curve of 0.747.

Discussion: GenoRisk could improve the risk assessment of individuals identified for prevention studies.

KEYWORDS
Alzheimer's disease heritability, Alzheimer's disease prevention studies, cross-validation, genetic risk, model validation, polygenic risk score, regression, risk models

1 | INTRODUCTION

Alzheimer's disease (AD) is notoriously difficult to treat, and the field has been plagued by an exceptionally high rate of clinical trial failures.[1–3] Part of the reason for these failures is the high heterogeneity across the disease that presents with diverse clinical symptoms and progression patterns.[4,5] Attempts to categorize various subtypes of AD have led to potentially more predictable outcomes.[6] Despite this, no disease-altering treatments exist.

1. Systematic Review: The authors reviewed the literature for Alzheimer's disease (AD) genetic markers and polygenic risk scores using traditional sources, meeting abstracts, and presentations. Recent clinical trials are considering inclusion of more than just apolipoprotein E ε4 genotype as a way of reducing variability in analysis of outcomes. Several polygenic risk scores are already available and were all constructed similarly.
2. Interpretation: By basing this model on individual data and incorporating genotype, age, and sex we were able to identify an

optimal-performing model that improves upon prior scores. These results were combined to develop a personalized genetic risk score, termed "age and sex-adjusted GenoRisk" that can be used for reducing variability in clinical outcome models as well as personal risk assessment.

3. Future Directions: This first version of GenoRisk was restricted to genetic variation that had been previously associated with AD. Future studies will incorporate genome-wide screens for variation as well as additional environmental factors.

A variety of risk factors can alter the likelihood of AD and disease course including sleep,[7] exercise,[8–10] nutrition,[8,9,11,12] sex, and age.[13] Moreover, AD heritability estimates range from 50% to 80%,[14–17] suggesting a large genetic component. The main genetic determinants of AD risk are variants in the gene apolipoprotein E (*APOE*), which account for one quarter of the heritability.[14,17] Areas under the curve (AUCs) associated with *APOE* alone range from 0.65 to 0.8 across different populations. It is likely the range in *APOE* association is due to additional genetic variation or environmental factors that differ across populations influencing AD onset and progression. Genome-wide association studies (GWAS) have identified more than 50 additional single nucleotide polymorphisms (SNPs) that contribute to the heritability of AD.[18,19,28,20–27] Identification of these genetic risk factors explains an additional 25% to 55% heritability and also highlights possible biological pathways of disease onset and progression.[17,29] Genetic risk scores that incorporate additional genetic variation to predict AD status range from an AUC of 0.57 (20 SNPs) to 0.8 (using > 200,000 SNPs including *APOE*). Despite the increase in genetic associations derived from the GWAS era, and thus increase in predictive ability afforded by genetic and lifestyle factors, most studies only control for *APOE* ε4 genotype (≈25% of heritability) in their predictive algorithms. Moreover, studies derive their risk scores from individual populations that may not accurately estimate risk.

Polygenic risk score (PRS) approaches that combine literature-derived odds ratios benefit from the large sample sizes that are usually available for the estimation of the odds ratios for each individual SNP. However, many recent PRS are based on and also validated against results from the International Genomics of Alzheimer's Project meta-analysis.[29–35] These PRS models rely on correlational assumptions that are difficult to verify and that do not frequently allow adjustment for patient-level factors such as sex and age. The phenotypic risk score using raw data has the advantage of simultaneous calculation of risk for correlated SNPs but is often based on a smaller sample size. However, because the model is built on raw data instead of meta-analysis data using odds ratios from each SNP considered for the model, the age effect can be directly estimated. Although case-control studies are limited because individuals may eventually develop AD, we accounted for this limitation by adjusting for the age of the individuals and weighted our estimates by population-based age distributions. Phenotypic prediction model approaches afford the most personalized approach because they are based on raw data.

In this article we used data from four independent genetic studies of AD to predict genetic risk. Moreover, we compared the efficacy of various phenotypic prediction models and used this information to create and validate a model comprised of 29 SNPs that predicts an individual's risk of developing AD, independent of age, sex, and genetic risk factors, termed the GenoRisk score. The GenoRisk model was found to be highly efficacious in its prediction of AD risk with an AUC of 0.747, a score that is comparable to PRS' reported in the literature with an equivalent number of SNPs[29] (AUC = 0.73, N = 28 SNPs) for early-onset AD. The age and sex–adjusted GenoRisk model further personalizes an individual's AD risk while detecting a broad range of risks within each *APOE* isoform allowing refinement in risk assessment for ε4 carriers and additional isoform combinations. Overall, the GenoRisk model is an accurate predictive model able to measure genetic risk within a population beyond the basic *APOE* genotypes.

2 | METHODS

2.1 | Study populations

Data consisting of 2579 cases and 2578 controls from the Alzheimer's Disease Genetics Consortium (ADGC) database were used to train and compare the various models tested for GenoRisk assessment. Datasets NG00026, NG00028, NG00030, NG00034, and NG00047 were used. Many study participants were White (84%), with 4% Black, and the remainder classified as "Other." Study population demographics are described elsewhere.[32]

The final GenoRisk model was calibrated using data from the 1000 Genomes Project participants to match the prevalence of SNPs in the general population; 1000 Genomes Project data were also used for GenoRisk score transformation.

2.2 | SNP selection and imputation

A total of 58 SNPs were selected through a literature search of all articles investigating the genetics of AD.[18,19,36,20,22–28] The SNPs considered are listed in Table S1 and Table S2 in supporting information. Because the ADGC data were collected on several different platforms, not all SNPs were available in all datasets, therefore the panel of SNPs listed in Table S1 were imputed using data from the 1000 Genomes Project. Imputation was performed within each population and subsequently combined as described in Ridge et al.[37]

2.3 | Statistics

2.3.1 | GenoRisk model fitting and selection

Genetic component: The combined imputed ADGC dataset consisting of 2579 cases and 2578 controls was used for GenoRisk model fitting and selection. An additive model was used for most SNPs tested apart from rs5848, rs4293518, and rs7412. A recessive model was used for rs5848.[38] *APOE* has three isoforms—ε2, ε3, and ε4—that are

characterized by varying combinations of two SNPs, rs4293518 and rs7412. Therefore, two different genetic models were tested for *APOE*:

1. Classification of *APOE* allelic variants (ε2, ε3, and ε4) using rs429358 and rs7412 genotypes and incorporating them into an allelic model;
2. Coding rs429358 and rs7412 into ε2, ε3, and ε4 isoforms and using a genotypic model for the six possible *APOE* genotypes.

Modeling: Creating a prediction score for a specific phenotype, such as AD, usually uses one of two broad approaches:

1. PRS: A risk score is created by combining odds ratios from the literature using an approach that assumes independence between SNPs. SNPs are either pruned[29] (only the SNP with the highest odds ratio out of a set of correlated SNPs is included in the analysis) or odds ratios from the literature are combined using a method that accounts for the correlations between SNPs such as those used by Stocker et al.[29] and Purcell et al.[39]
2. Phenotypic prediction model: A risk score is created using a phenotypic prediction model (i.e., linear, logistic, lasso, or elastic net regression) based on an original dataset with individual-level phenotypes and SNPs. Odds ratios for each SNP are calculated within the model and simultaneously account for correlations. Historically, this type of approach has been impractical due to the large data sets and high computational requirements, but these methods are now more practical with the availability of more powerful computing resources.

To this end, four general statistical methods were tested: logistic, probit, lasso, and elastic net regression. Lasso and elastic net were based on logistic regression. The elastic net initially used $\lambda = 0.5$ but was refined through cross-validation. All statistical models included a genetic component along with age and sex. Models were tested with and without the age × sex interaction term.

In addition to the standard genetic models described above, a genetic model that incorporated the odds ratios estimated from previous studies was tested. In this class of models, all the estimates for the odds ratios from previous studies were combined into a single score for each individual, and that score was included in the model as a covariate with age, sex, and *APOE* status. The scores were either the sum of the effects of the SNPs (additive) or the product (multiplicative) of the effects of the SNPs.

Finally, linkage disequilibrium (LD), which is a correlation between SNPs that tends to increase for SNPs that are physically closer to each other on the chromosome, can lead to overfitting when combining estimates that were calculated independently. To account for this, we used LD pruning, which is a method of eliminating lower risk SNPs that are correlated with higher risk SNPs, to create an LD-pruned multiplicative score and an LD-pruned additive score. LD pruning was performed by calculating the LD between each pair of SNPs that shared a chromosome. When two SNPs had $r^2 > 0.2$, the SNP with the lower effect as estimated from the literature was dropped.

2.3.2 | GenoRisk model validation

Models were compared using 40 repetitions of 10-fold cross-validation. The Brier score was used to determine the accuracy of the predictions (lower is more accurate). In each repetition, the mean of the 10 repetitions was calculated, then summaries of the 40 repetitions were compared across the different models. Because lasso regression is a special case of the elastic net (lasso is the elastic net when $\lambda = 0$), the elastic net model was selected, then the selection of λ was further optimized through cross-validation until the mean Brier score was minimized. To further validate the genotypic estimates from the best performing (final GenoRisk) model, the *APOE* estimates were compared to the estimates obtained in Genin et al.[40]

2.3.3 | GenoRisk score transformation

Because the GenoRisk output is a proportion that could be incorrectly interpreted as percent risk, the GenoRisk output was transformed to fit on a scale from 0 to 40 based on a the max and min GenoRisk scores observed in the 1000 Genomes Project. To calculate GenoRisk, call γ the genetic portion of the logit curve, so $\gamma = \sum_{i=1}^{32} \beta_i x_i$, where each x_i represents either one of the 29 SNPs or one of the three *APOE* isoforms, and each β_i is the coefficient for that genetic risk factor estimated from the GenoRisk model, and let Γ_μ be the set of logit scores for 1000 Genomes Project subjects based on their genetic risk, γ, and an intercept μ. Define μ such that $40\frac{\{[1+exp(-\mu-\gamma)]^{-1}-min(\Gamma_\mu)\}}{(\Gamma_\mu)}$ is 20 when the probability of having AD at age 85 is 10%.

2.3.4 | Absolute AD risk calculation

Unconditional: The Silverman dataset[41] was used to derive age, sex, and model intercept estimates that are more representative of a normal population. A preliminary logit curve was fit to the Silverman results. The curve was assumed to represent the quantile median (ε3/ε3 genotype) and to be balanced between males and females. The model also assumes that the data are representative of median risk from the other 29 SNPs in the GenoRisk model. These Silverman age and intercept estimates were combined with the GenoRisk genotypic and sex estimates above to make the final absolute AD algorithm, which outputs an estimated probability of developing AD.

Conditional: Assuming that a person's unconditional risk is r_0, then his or her risk conditional on not-currently-having-AD is:

$$f\ (x|r_0) = \frac{f\,(x) - r_0}{1 - r_0}$$

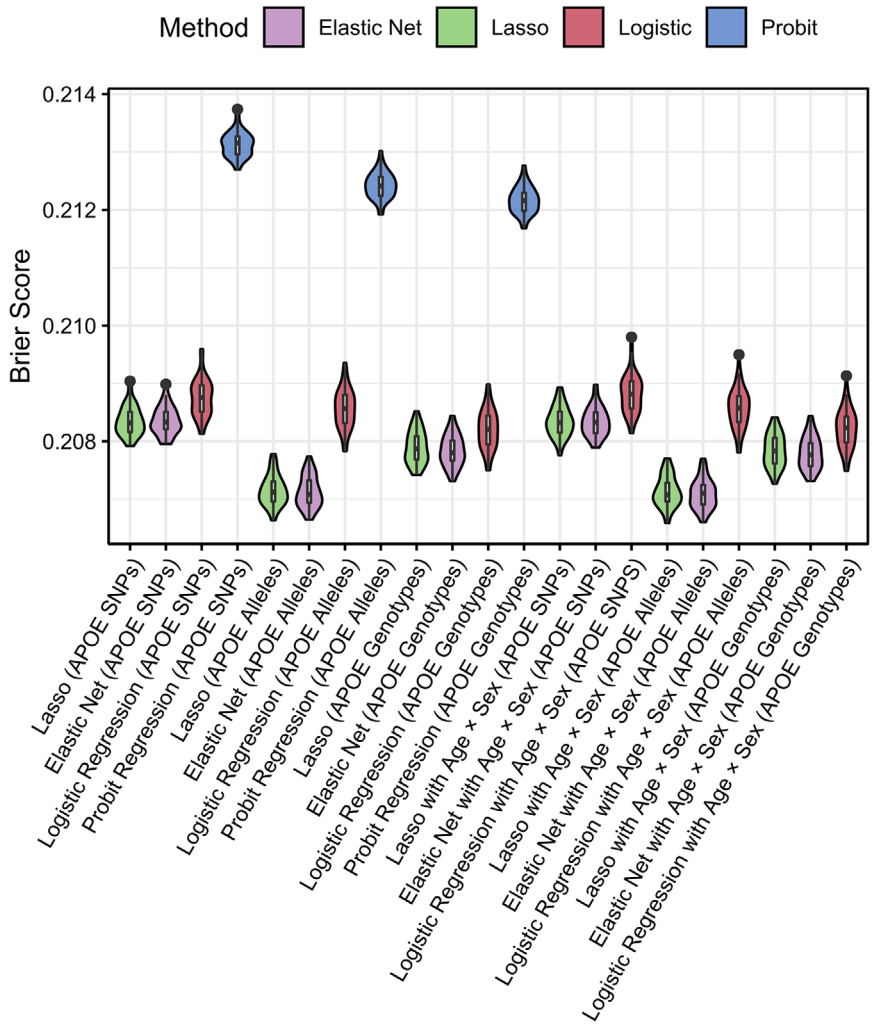

FIGURE 1 Violin plots comparing the Brier scores derived from 10-fold cross-validation for 21 of the 25 models tested. Elastic net, lasso, logistic, and probit are shown in purple, green, red, and blue, respectively. The scores for the other four models are not shown here because their Brier scores were so much higher than the other models that it altered the scale of the figure and made comparison between the remaining models more difficult

3 | RESULTS

3.1 | Derivation of GenoRisk model

The summaries of the Brier scores from the 40 repetitions of cross-validation on the 25 models are presented in **Table 1 and in Figure 1**. The eight highest performing models were the regularization methods (lasso and elastic net) while the top four models were regularization methods with an allelic model for *APOE*. The top two models used the elastic net, and the best model based on the mean Brier score was the allelic *APOE* elastic net model that included the age × sex interaction (mean Brier score = 0.20709; AUC = 0.747) and is used to calculate an individual's GenoRisk score. Because elastic net allows variables to drop out of the model, the final model included the intercept, the three *APOE* isoforms, 29 other SNPs, age, and age × sex interaction. The distribution of GenoRisk scores stratified by isoform status are shown in Figure 2.

To further validate GenoRisk model performance, the estimated odds ratios for *APOE* isoforms from the GenoRisk score were compared to *APOE* isoform odds ratios reported in Genin et al.,[40] which were derived from 7531 cases and 10,132 controls. The estimated odds ratios from this final GenoRisk model fell within the confidence intervals for the odds ratios estimated by Genin et al.[40] for all four isoforms (Figure 3), indicating that the additional SNPs did not alter model estimation of risk due to *APOE* and likely overall model estimates.

3.2 | Derivation of absolute AD risk estimates

The GenoRisk model is independent of age and sex; however, age and sex effects are known to alter absolute risk for various diseases including AD. While the ADGC data are useful for estimating coefficients for the genetic risk factors and are appropriate to make estimates for differences between sexes, only an appropriately designed prospective study can accurately estimate the age-related incidence of AD. To finalize the model and be able to provide accurate estimates of the age-related risk of developing AD based on the GenoRisk score, we used estimates of cumulative risk of developing AD from Silverman et al.[41] A logit curve was fit to the dataset that excluded parents and siblings of individuals with early onset AD and very late onset AD. This is expected to give a more accurate representation of the risk of developing AD in the general population as it excludes individuals known to be biased

TABLE 1 A summary of the mean Brier score from each of the 40 repetitions of the 10-fold cross-validation training-and-test trials for each of the 25 models. Models are ordered from most accurate to least according to the 40 mean Brier scores

Method	*APOE* model	Age × sex interaction	Min.	1st quartile	Median	Mean	3rd quartile	Max.
Elastic net	Allelic	Yes	0.20661	0.20691	0.20710	0.20709	0.20725	0.20770
Elastic net	Allelic	No	0.20665	0.20694	0.20708	0.20713	0.20733	0.20774
Lasso	Allelic	Yes	0.20659	0.20696	0.20707	0.20714	0.20728	0.20770
Lasso	Allelic	No	0.20664	0.20696	0.20712	0.20715	0.20730	0.20778
Elastic net	Genotypic	Yes	0.20731	0.20757	0.20777	0.20780	0.20797	0.20843
Elastic net	Genotypic	No	0.20731	0.20767	0.20780	0.20784	0.20801	0.20844
Lasso	Genotypic	Yes	0.20727	0.20762	0.20783	0.20785	0.20806	0.20841
Lasso	Genotypic	No	0.20742	0.20769	0.20787	0.20790	0.20809	0.20852
Logistic regression	Genotypic	No	0.20750	0.20795	0.20821	0.20820	0.20844	0.20899
Logistic regression	Genotypic	Yes	0.20749	0.20798	0.20823	0.20822	0.20843	0.20913
Lasso	SNP	Yes	0.20776	0.20815	0.20832	0.20834	0.20852	0.20893
Elastic net	SNP	Yes	0.20789	0.20814	0.20833	0.20834	0.20850	0.20898
Elastic net	SNP	No	0.20796	0.20821	0.20833	0.20835	0.20851	0.20899
Lasso	SNP	No	0.20792	0.20816	0.20832	0.20836	0.20851	0.20904
Logistic regression	Allelic	No	0.20783	0.20832	0.20856	0.20856	0.20881	0.20936
Logistic regression	Allelic	Yes	0.20781	0.20833	0.20857	0.20857	0.20878	0.20950
Logistic regression	SNP	No	0.20813	0.20851	0.20876	0.20875	0.20897	0.20959
Logistic regression	SNP	Yes	0.20814	0.20856	0.20882	0.20882	0.20904	0.20980
Probit regression	Genotypic	No	0.21168	0.21198	0.21216	0.21217	0.21230	0.21276
Probit regression	Allelic	No	0.21192	0.21225	0.21241	0.21242	0.21257	0.21301
Probit regression	SNP	No	0.21270	0.21296	0.21316	0.21313	0.21327	0.21374
Pruned multiplicative score	SNP	No	0.22277	0.22291	0.22296	0.22296	0.22302	0.22310
Pruned additive score	SNP	No	0.22541	0.22559	0.22566	0.22565	0.22571	0.22578
Full multiplicative score	SNP	No	0.22713	0.22725	0.22730	0.22730	0.22736	0.22747
Full additive score	SNP	No	0.23607	0.23624	0.23634	0.23632	0.23640	0.23644

Abbreviations: *APOE*, apolipoprotein E; SNP, single nucleotide polymorphism.

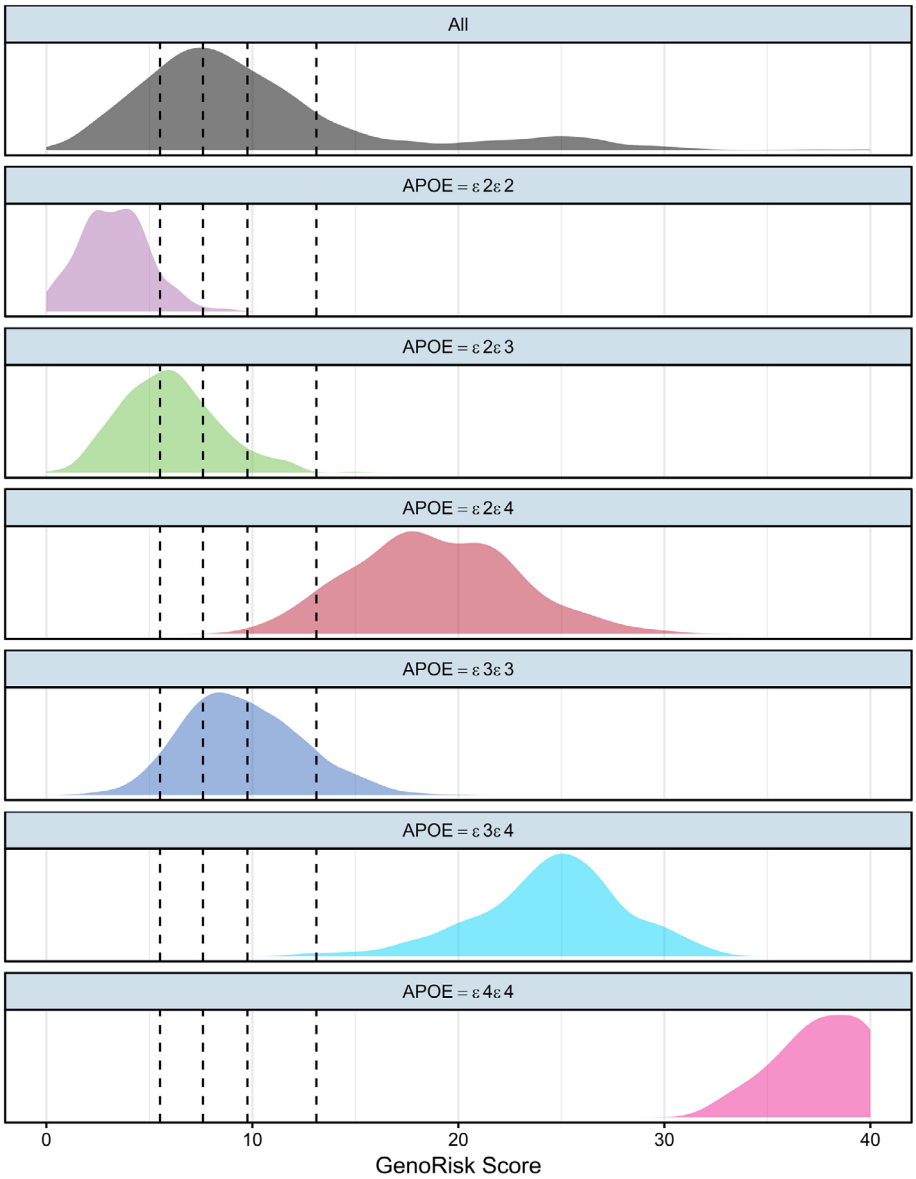

FIGURE 2 The distribution of GenoRisk scores among the 2504 individuals in the 1000 Genomes Project. The scores are defined so that a score of 20 indicates a 10% risk of having Alzheimer's disease at the age of 85. The 40-point scale helps reduce the risk of misinterpreting the score as a probability. The vertical dashed lines divide the overall distribution into quintiles

heavily by genotypic effects. The final adjusted model plotted over the Silverman et al. data is shown in Figure S1 in supporting information. The logit curve was fit using the data from individuals ≤ 85 years of age. Incorporation of individuals over 85 years resulted in a less reliable model as this subpopulation appears to have a reduced rate of risk accrual over time; this is likely due to competing risk within this age group.[42] Estimands (age and intercept) from the logit fit were incorporated with GenoRisk genotypic and sex estimands to generate age and sex adjusted GenoRisk estimates appropriate for the general population.

Age- and sex-specific risk given a subject who does not currently have AD (unconditional risk) may also be calculated. An example is presented in Figure 4.

4 | DISCUSSION

We developed a genetic prediction model of AD, termed GenoRisk Score. While others have derived PRSs, our assessment of several different methods compared through cross-validation coupled with genetic estimates derived across multiple populations with varying ethnic backgrounds and use of a more appropriate population to estimate age-specific risk of having AD make this work unique. This work also benefits from the ability of the elastic net methodology to identify models in the presence of correlated endpoints. Compared to PRSs in the literature which require a priori omission of correlated SNPs or other post hoc adjustments to account for correlation, the GenoRisk score reduces the risk of overfitting by incorporating patient-level data

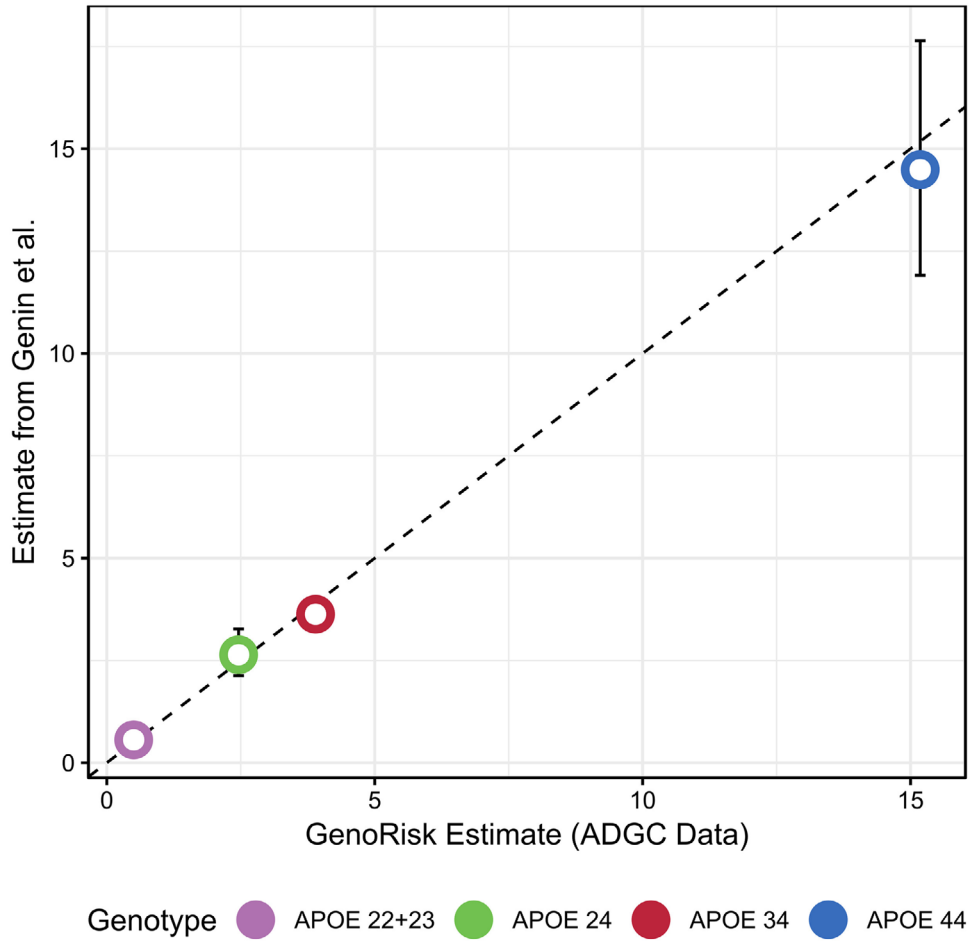

FIGURE 3 The estimated odds ratios of the apolipoprotein E genotypes from the final GenoRisk model compared to the estimates from Genin et al.[40] To combine ε2/ε2 with ε2/ε3 in GenoRisk to match the format of Genin et al. the effect shown is the weighted mean of the effect of the two genotypes. The ε3/ε3 genotype was the reference genotype in the Genin et al. comparison and ε3 was the reference allele in the GenoRisk model, so the odds ratios for both are 1, by definition. The dashed line indicates the line at which the two estimates are equal

using external cross validation. Furthermore, consistency in odds ratios for *APOE* isoform estimation calculated from our model and the literature, provides outside validation of our results and confirms that our model appropriately estimates *APOE* effects even in the presence of several other genes.

Moreover, AUC from PRS ranges are reported from 0.57 to 0.84; this range includes studies of early onset AD which has a larger genetic component (AUC 0.73 with 28 SNPs), as well as PRS' generated using thousands of SNPs (AUC 0.75–0.84; > 4000 SNPs). PRS' generated with larger numbers of SNPs, although improving in processing time and cost, are not broadly feasible for clinical assessment. The GenoRisk platform incorporates a smaller number of SNPs to estimate AD risk. Although the GenoRisk AUC is marginally less predictive than models incorporating thousands of SNPs (AUC 0.747), this model performs better than previously published PRS' of equivalent size with a similar AD population (0.63 and 0.70).[29]

Although *APOE* ε4 is the single greatest genetic determinant of AD risk (apart from rare mutations in *PS1*, *PS2*, and *APP* responsible for autosomal-dominant familial AD), there is a relatively wide range of risk within a given *APOE* genotype, as shown in this and other studies.[29] In some cases, individuals with a low-risk *APOE* genotype (e.g., ε3/ε3) may be revealed to have a higher overall genetic risk than some patients with a high-risk *APOE* genotype (e.g., ε3/ε4) when other genetic variants are accounted for, and vice versa. Currently, many trials use *APOE* ε4 status alone to segregate participants into high-risk and low-risk categories, which may reduce study power. The utility of GenoRisk for assessing personalized risk can be demonstrated using an example of a 72-year-old male who wants to know his risk of developing AD. Initial assessment by sex only indicates that his risk will follow the population average for males (Figure 5; dashed blue line). Incorporation of individual *APOE* genotype (ε3/ε4; most common genetic factor incorporated in clinical trials), increases absolute AD risk well above the population average for males (dashed black line). Calculation of the absolute AD risk via incorporation of multiple genetic drivers of AD through the GenoRisk model (dashed green line) reduces AD risk compared to *APOE* status alone. Finally, risk adjustment for individual current AD status (conditional assessment; solid green line) further reduces predicted AD risk for this individual, that is, given that he does not currently have AD at his present age reduces his overall likelihood of developing it in the future. Personalized genetic risk classification,

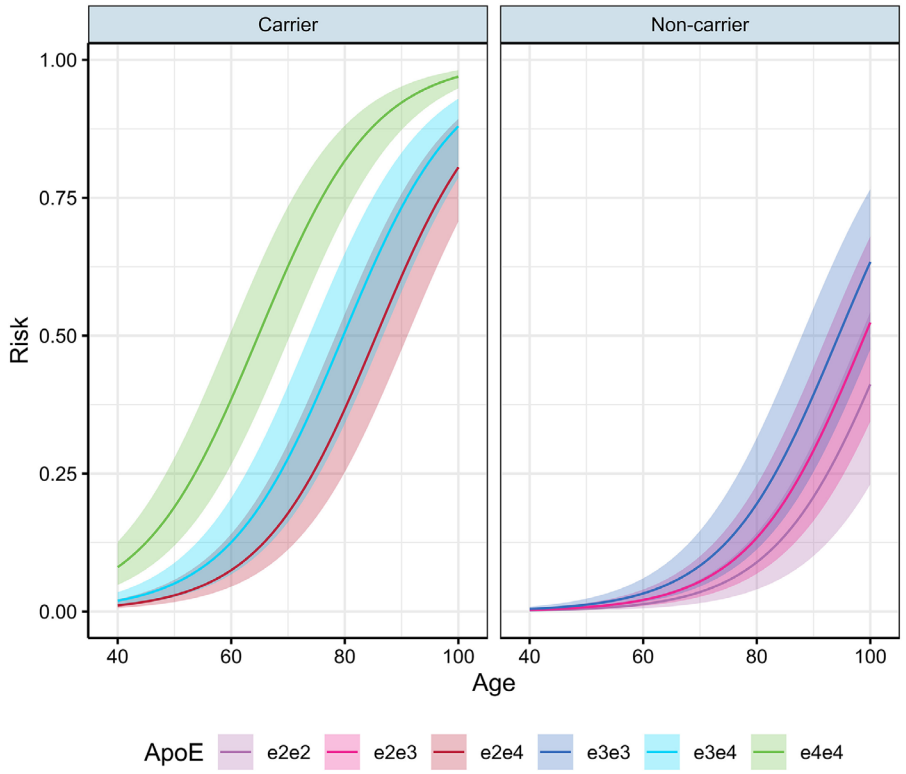

FIGURE 4 Lineplots of 1000 Genomes Project age and sex-adjusted GenoRisk probabilities stratified by age and apolipoprotein E isoform status. The solid line represents the mean value of individuals within each isoform group (+ 95% confidence intervals) across the age ranges present

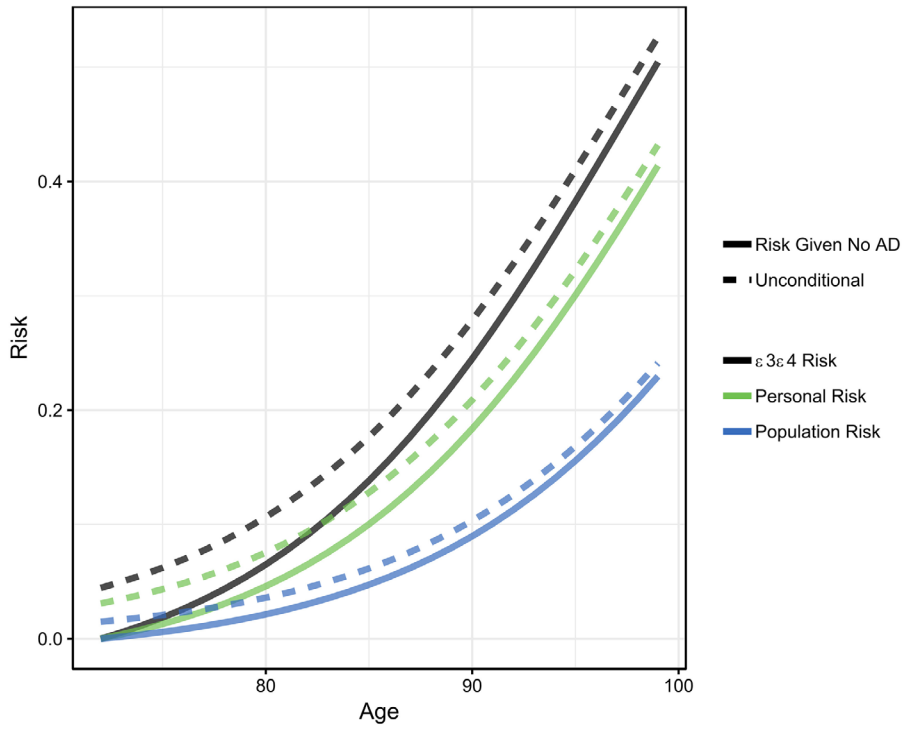

FIGURE 5 Probability by age of having Alzheimer's disease (AD; unconditional risk), or developing AD given that it is not present (solid curves line), in a hypothetical 72-year-old male. Comparison of curves for three conditions: given no additional information (blue), given that the individual has the ε3/ε4 genotype (black), and given the additional information provided by the GenoRisk score (green). Specific risk curves can help provide better metrics with which to compare treatment outcomes in future clinical trials

afforded by GenoRisk, may improve prediction of decline rate in clinical trials, improving power for detecting treatment differences as well as aid in identification of personalized treatments or preventative measures.

Limitations of our GenoRisk score are primarily a result of having a relatively small sample size for building the model. This small sample size required us to assume the absence of age by genotype interactions, which is likely to be an inaccurate assumption. The risk model relies on the diagnosis, which is purely clinical, and not biomarker based, and the misclassification may be as high as 10% with experienced clinicians, but potentially closer to 30% with less experienced clinicians. Validation of the estimation of age-related risk and corresponding timing of diagnosis would require a prospective, longitudinal study. In addition, our population has primarily European ancestry, limiting the generalizability to other populations. The analysis relies on imputation due to missing genotypes, reducing the accuracy of the model.

Future studies will be aimed at assessing whether subcategorization of individuals based on specific genetic variants or sets of genetic variants are associated with specific neuropathologic subtypes or clinical presentation patterns. These subcategories of patients may respond differently to treatments, may progress at different rates, or may differ in biomarker patterns associated with disease progression. In general, GenoRisk provides a better understanding of personalized AD risk that could facilitate the development of new treatments or preventive measures.

ACKNOWLEDGMENTS

This research was funded by AFFIRMATIV diagnostics.

CONFLICTS OF INTEREST

Suzanne Hendrix, as the owner of Pentara, received support from Affirmative Diagnostics for the manuscript and received contract work from AC Immmune, Acumen, ADCS, ADDF, Affirmativ Dignpostics, Alector, Alkahest, Allergan, Alzheon, Amylyx, Apodemus, Athira, Avanir, Banner, Biogen, Biohaven, Cadent, Capricor, Cerecin, Cognito, Cognoptix, Cortexyme, Elusis, Green Valley, Grifols, Ionis, Janssen, Kyowa-Kirin, Lexeo, LuMind, Lundbeck, NCRI, Nilvad, NovoNordisk, Nutricia, Photopharmics, PTCBio, Regenera, Retrotope, Richard Isaacson, Samus, Soleno, Spaulding, Suven, Takeda, Telocyte, Tetra Discovery, Toyama, United Neuroscience, Vaccinex, Vacinity, vTv Therapeutics, and WGU. Suzanne Hendrix has also received royalties from Wiley in the last 36 months; received honoraria for lectures, presentations, etc. from Biogen; received support in the last 36 months from Grifols and Nutricia for presentations and conference attendance; claims one patent pending Brown BL, Hedges DW, Hendrix SB (2014) Extracting A periodic Components from a Time Series Wave Data Set—Patent Pending with a filing date of October 16, 2013, Application Number PCT/US2013/065327; and participated in Alzheon, Cortexyme, and Janssen DSMB or Advisory Board in the last 36 months. Samuel P. Dickson received support from Pentara, which in turn received support from Affirmativ Diagnostics for the manuscript. S. B. Booth received support from Affirmative Diagnostics for the manuscript; received travel support from her employer for work events; ADX provides profit sharing incentives with additional salary which were broadly applied to any work performed for the company (including manuscript generation). J. S. K. Kauwe received support for the manuscript from Brigham Young university and has received honoraria from UH Hilo. B. L. Brown received support through the following grants: Measuring the Interactive Effects of COVID-19 and Latent Infections on Patterns of cognitive Timing: A Middle-Aged to Elderly Sample from the Utah Valley Community. Translational Medicine Award from the BYU Simmons Research Endowment ($31,300). This is an intramural grant from a research endowment within our university to my departmental research fund. Brown also claims one patent pending Brown BL, Hedges DW, Hendrix SB (2014) Extracting A periodic Components from a Time Series Wave Data Set—Patent Pending with a filing date of October 16, 2013, Application Number PCT/US2013/065327. Data used in this manuscript is publically available. Participant consent is not needed, and was obtained from investigators who ran each study. M. L. Hardy has received paid consulting fees to the LLC solely owned by ML Hardy from ixLayer, NW Pathology & Labs, Mt. Baker Imaging, PlumCare, ADx Healthcare, Architectural Elements; received travel support from her employer for work events; and was the past Volunteer Secretary for the Lets Pool Together and Hope Philantropies non-profit boards. P. G. Ridge had received honoraria from University of Kansas Alzheimer's Disease Center for a seminar presentation. R. R. Fortna is a board member of Northwest Pathology and ADx Healthcare and is co-owner and holds stock for ADx healthcare. J. Nicodemus Johnson has nothing to disclose.

REFERENCES

1. Cummings J, Feldman HH, Scheltens P. The "rights" of precision drug development for Alzheimer's disease. *Alzheimer's. Res Ther*. 2019;11:1-14
2. Cummings J, Lee G, Ritter A, Zhong K. Alzheimer's disease drug development pipeline: 2018. *Alzheimer's Dement Transl Res Clin Interv*. 2018;4:195-214.
3. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. Alzheimer's. *Res Ther*. 2014;6:37.
4. Chaudhury S, Brookes KJ, Patel T, et al. Alzheimer's disease polygenic risk score as a predictor of conversion from mild-cognitive impairment. *Transl Psychiatry*. 2019;9:1-7.
5. Neu SC, Pa J, Kukull W, et al. Apolipoprotein E genotype and sex risk factors for Alzheimer's disease: a meta-analysis. *JAMA Neurol*. 2017;74:1178-1189.
6. Murray ME, Graff-Radford NR, Ross OA, Petersen RC, Duara R, Dickson DW. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *Lancet Neurol*. 2011;10:785-796.
7. Lim ASP, Kowgier M, Yu L, Buchman AS, Bennett DA Sleep fragmentation and the risk of incident Alzheimer's disease and cognitive decline in older persons. *Sleep*. 2013;36:1027-1032..
8. Clare L, Wu Y-T, Teale JC, et al. Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: A cross-sectional study 2017.
9. Gorelick PB, Furie KL, Iadecola C, et al. Defining optimal brain health in adults: a presidential advisory from the American Heart Association/American Stroke Association. *Stroke*. 2017;48.

10. Smith PJ, Blumenthal JA, Hoffman BM, et al. Aerobic exercise and neurocognitive performance: a meta-analytic review of randomized controlled trials. *Psychosom Med*. 2010;72:239-252.
11. Singh B, Parsaik AK, Mielke MM, et al. Association of mediterranean diet with mild cognitive impairment and Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimer's Dis*. 2014;39:271-282.
12. Morris MC, Tangney CC, Wang Y, et al. MIND diet associated with reduced incidence of Alzheimer's disease. *Alzheimer's Dement*. 2015;11:1007-1014.
13. Mielke MM. Sex and gender differences in Alzheimer's disease dementia. *Psychiatr Times*. 2018;35:14-17.
14. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer's disease: clinical implications and perspectives. *Genet Med*. 2016;18:421-430.
15. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer's disease. *Arch Gen Psychiatry*. 2006;63:168.
16. Pedersen NL, Gatz M, Berg S, Johansson B. How heritable is Alzheimer's disease late in life? Findings from Swedish twins. *Ann Neurol*. 2004;55:180-185.
17. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK Alzheimer's disease: analyzing the missing heritability. *PLoS One*. 2013;8:e79771.
18. Escott-Price V, Bellenguez C, Wang L-S, et al. Gene-Wide analysis detects two new susceptibility genes for Alzheimer's disease. *PLoS One*. 2014;9:e94661.
19. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45:1452-1458.
20. Wetzel-Smith MK, Hunkapiller J, Bhangale TR, et al. A rare mutation in UNC5C predisposes to late-onset Alzheimer's disease and increases neuronal cell death. *Nat Med*. 2014;20:1452-1457.
21. Carrasquillo MM, Belbin O, Hunter TA, et al. Replication of CLU, CR1, and PICALM associations with Alzheimer's disease. *Arch Neurol*. 2010;67:961-964.
22. Guerreiro R, Wojtas A, Bras J, et al. TREM2 variants in Alzheimer's disease. *N Engl J Med*. 2013;368:117-127.
23. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*. 2009;41:1088-1093.
24. Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet*. 2011;43:429-435.
25. Jun G, Naj AC, Beecham GW, et al. Meta-analysis Confirms CR1, CLU, and PICALM as Alzheimer's disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol*. 2010;67:1473.
26. Kamboh MI, Demirci FY, Wang X, et al. Genome-wide association study of Alzheimer's disease. *Transl Psychiatry*. 2012;2:e117-e117.
27. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet*. 2011;43:436-441.
28. Ridge PG, Karch CM, Hsu S, et al. Linkage, whole genome sequence, and biological data implicate variants in RAB10 in Alzheimer's disease resilience. *Genome Med*. 2017;9:100.
29. Stocker H, Möllers T, Perna L, Brenner H. The genetic risk of Alzheimer's disease beyond APOE ε4: systematic review of Alzheimer's genetic risk scores. *Transl Psychiatry*. 2018;8:166.
30. Escott-Price V, Sims R, Bannister C, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain*. 2015;138:3673-3684.
31. Lee SH, Harold D, Nyholt DR, et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum Mol Genet*. 2013;22:832-841.
32. Desikan RS, Fan CC, Wang Y, et al. Genetic assessment of age-associated Alzheimer's disease risk: development and validation of a polygenic hazard score. *PLOS Med*. 2017;14:e1002258.
33. Marden JR, Mayeda ER, Walter S, Vivot A. Using an Alzheimer's disease polygenic risk score to predict memory decline in black and white Americans over 14 years of follow-up. *Alzheimer's Dis Assoc Disord*. 2016;30:195-202.
34. Escott-Price V, Shoai M, Pither R, Williams J, Hardy J. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. *Neurobiol Aging*. 2017;49:214.e7-214.e11.
35. Mormino EC, Sperling RA, Holmes AJ, et al. Polygenic risk of Alzheimer's disease is associated with early- and late-life processes. *Neurology*. 2016;87:481-488.
36. Carrasquillo MM, Belbin O, Hunter TA, et al. Replication of CLU, CR1, and PICALM associations with Alzheimer's disease. *Arch Neurol*. 2010;67:961-964.
37. Ridge PG, Hoyt KB, Boehme K, et al. Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol Aging*. 2016;41:200.e13-200.e20.
38. Sheng J, Su L, Xu Z, Chen G. Progranulin polymorphism rs5848 is associated with increased risk of Alzheimer's disease. *Gene*. 2014;542:141-145..
39. International Schizophrenia Consortium S, Purcell SM, Wray NR, Stone JL, et al, International Schizophrenia Consortium S. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748-752.
40. Genin E, Hannequin D, Wallon D, et al. APOE and Alzheimer's disease: a major gene with semi-dominant inheritance. *Mol Psychiatry*. 2011;16:903-907.
41. Silverman JM, Smith CJ, Marin DB, Mohs RC, Propper CB. Familial patterns of risk in very late-onset Alzheimer's disease. *Arch Gen Psychiatry*. 2003;60:190..
42. Chang C-CH, Zhao Y, Lee C-W, Ganguli M. Smoking, death, and Alzheimer's disease: a case of competing risks. *Alzheimer's Dis Assoc Disord*. 2012;26:300-306.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dickson SP, Hendrix SB, Brown BL, et al. GenoRisk: A polygenic risk score for Alzheimer's disease. *Alzheimer's Dement*. 2021;7:e12211. https://doi.org/10.1002/trc2.12211