



Members of the EADB, GR@ACE, DEGESCO, DemGene, GERAD, and EADI Groups (2022). Association of Rare APOE Missense Variants V236E and R251G With Risk of Alzheimer Disease. *JAMA Neurology*, 79(7), 652-663.
<https://doi.org/10.1001/jamaneurol.2022.1166>

Peer reviewed version

License (if available):
CC BY

Link to published version (if available):
[10.1001/jamaneurol.2022.1166](https://doi.org/10.1001/jamaneurol.2022.1166)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via American Medical Association at <https://doi.org/10.1001/jamaneurol.2022.1166>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Rare *APOE* missense variant R251G is associated with reduced risk of Alzheimer's disease in *APOE-ε4* carriers

Authors: Yann Le Guen , PhD ^{1,2,*,#}, Michael E. Belloy , PhD ^{1,*}, Benjamin Grenier-Boley , MSc ³, Itziar de Rojas , MSc ^{4,5}, Atahualpa Castillo-Morales , PhD ⁶, Iris Jansen , PhD ^{7,8}, Aude Nicolas , PhD ³, Céline Bellenguez , PhD ³, Carolina Dalmaso , PhD ^{9,10}, Fahri Küçükali , PhD ^{11,12,13}, Sarah J. Eger ¹, Katrine Laura Rasmussen , MD,PhD ^{14,15}, Jesper Qvist Thomassen , PhD ¹⁴, Jean-François Deleuze , PhD ¹⁶, Zihuai He , PhD ^{1,17}, Valerio Napolioni , PhD ¹⁸, Philippe Amouyel , MD,PhD ¹⁹, Frank Jessen , MD ^{20,21,22}, Patrick G. Kehoe , PhD ²³, Cornelia Van Duin , PhD ^{24,25}, Magda Tsolaki , MD,PhD ²⁶, Pascual Sánchez-Juan , MD,PhD ^{27,5}, Kristel Slegers , MD,PhD ^{11,12,13}, Martin Ingelsson , MD,PhD ^{28,29,30}, Giacomina Rossi , PhD ³¹, Mikko Hiltunen , PhD ³², Rebecca Sims , PhD ³³, Wiesje M. van der Flier , PhD ⁷, Alfredo Ramirez , MD,PhD ^{9,34,35,21,22}, Ole A. Andreassen , MD,PhD ^{36,37}, Ruth Frikke-Schmidt , MD,PhD ^{14,15}, Julie Williams , PhD ^{6,33}, Agustín Ruiz , MD,PhD ^{4,5}, Jean-Charles Lambert , PhD ³, Michael D. Greicius , MD ¹, and the EADB group†, the GR@ACE group†, the DEGESCO group†, the EADI group†, the GERAD group†, for the EADB collaborator†, the GR@ACE collaborator†, the DEGESCO collaborator†, the DemGene collaborator†, the GERAD collaborators†.

* Co-first authors

†Data used in the preparation of this article were obtained from the EABD, GR@ACE, DEGESCO, DemGene, EADI, GERAD. As such, the investigators within these consortia contributed to the design and implementation of their respective consortium and/or provided data. Group authors who participated in analysis or writing of this report are listed at the end of this article. A full list of collaborators, including those who did not participate in analysis or writing of this report, is provided in the supplementary materials.

Correspondence should be addressed to:

Yann Le Guen

Department of Neurology and Neurological Sciences – Greicius lab

Stanford University

290 Jane Stanford Way, E265, CA 94305-5090

Tel: 650 666 2696

Email: yleguen@stanford.edu

Word count: 3257 words

1 **Group authors** (*authors may be listed in two groups, authors already in the byline are not listed*)

2
3 **EABD group authors**

4 Beatrice Arosio, PhD 38, Luisa Benussi, PhD 39, Anne Boland, PharmD, PhD 16, Barbara Borroni,
5 MD 40, Paolo Caffarra, PhD 41, Delphine Daian 16, Antonio Daniele, MD, PhD 42,43, Stéphanie
6 Debette, MD, PhD 44,45, Martin Dichgans, MD 46,47,48, Carole Dufouil, PhD 49,50, Emrah
7 Düzel, MD, PhD 51,52, Daniela Galimberti, PhD 53,54, Vilmantas Giedraitis, PhD 28, Timo
8 Grimmer, PhD 55, Caroline Graff, MD 56, Edna Grünblatt, PhD 57,58,59, Olivier Hanon, MD 60,
9 Lucrezia Hausner, PhD 61, Stefanie Heilmann-Heimbach, PhD 62, Henne Holstege, PhD 7,63,
10 Jakub Hort, MD, PhD 64,65, Deckert Jürgen, MD 66, Teemu Kuulasmaa 32, Aad van der Lugt,
11 PhD 67, Carlo Masullo, MD 68, Patrizia Mecocci, MD, PhD 69, Shima Mehrabian, PhD 70,
12 Alexandre de Mendonça, PhD 71, Susanne Moebus, PhD 72, Benedetta Nacmias, PhD 73,74, Gael
13 Nicolas, MD, PhD 75, Robert Olaso, PhD 16, Goran Papenberg, PhD 76, Lucilla Parnetti, MD, PhD
14 77, Florence Pasquier, MD, PhD 78, Oliver Peters, MD 79,80, Yolande A.L. Pijnenburg, MD 7,
15 Julius Popp, MD 81,82,83, Innocenzo Rainero, MD, PhD 84, Inez Ramakers, PhD 85, Steffi
16 Riedel-Heller, MD 86, Dan Rujescu, MD 87, Nikolaos Scarmeas, MD, PhD 88,89, Philip
17 Scheltens, MD, PhD 7, Norbert Scherbaum, MD 90, Anja Schneider, MD 21,91, Davide Seripa,
18 PhD 92, Hilikka Soininen, MD, PhD 93, Vincenzo Solfrizzi, MD, PhD 94, Gianfranco Spalletta,
19 MD, PhD 95,96, Alessio Squassina, PhD 97, John van Swieten, MD 98, Thomas J Tegos, MD, PhD
20 26, Lucio Tremolizzo, MD, PhD 99, Frans Verhey, MD, PhD 100, Martin Vyhnalek, MD, PhD
21 64,65, Jens Wiltfang, MD 101,102,103

22
23 **GRA@CE group authors**

24 Mercè Boada, MD, PhD 4,5, Pablo García-González, MSc 4,5, Raquel Puerta, MSc 4, Luis M
25 Real, PhD 104,105

26
27 **DEGESCO group authors**

28 Victoria Álvarez, PhD 106,107, María J. Bullido, PhD 108,5,109, Jordi Clarimon, PhD 110,5, José
29 María García-Alberca, MD, PhD 111,5, Pablo Mir, MD, PhD 112,5, Fermin Moreno, MD, PhD
30 113,5,114, Pau Pastor, MD, PhD 115,116, Gerard Piñol-Ripoll, MD, PhD 117,118, Laura Molina-
31 Porcel, MD, PhD 119,120, Jordi Pérez-Tur, PhD 121,5,122, Eloy Rodriguez-Rodriguez, MD, PhD
32 123,5, Jose Luis Royo, PhD 124, Raquel Sánchez-Valle, MD, PhD 125

33
34 **EADI group authors**

35 Carole Dufouil, PhD 49,50, Olivier Hanon, MD 60, Florence Pasquier, MD, PhD 78

36
37 **GERAD authors**

38 Martin Dichgans, MD 46,47,48, Dan Rujescu, MD 87

Affiliations

1. Department of Neurology and Neurological Sciences, Stanford University, Palo Alto, California, United States of America.
2. Institut du Cerveau - Paris Brain Institute - ICM, Paris, France.
3. Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1167-RID-AGE Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, F-59000 Lille, France.
4. Research Center and Memory clinic Fundació ACE, Institut Català de Neurociències Aplicades, Universitat Internacional de Catalunya, Barcelona, Spain.
5. Networking Research Center on Neurodegenerative Diseases (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain.
6. UKDRI@Cardiff, School of Medicine, Cardiff University, Wales, United Kingdom.
7. Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.
8. Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije University, Amsterdam, The Netherlands.
9. Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany.
10. Estudios en Neurociencias y Sistemas Complejos (ENyS) CONICET-HEC-UNAJ.
11. Complex Genetics of Alzheimer's Disease Group, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium.
12. Laboratory of Neurogenetics, Institute Born - Bunge, Antwerp, Belgium.
13. Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium.
14. Department of Clinical Biochemistry, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark.
15. Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.
16. Université Paris-Saclay, CEA, Centre National de Recherche en Génomique Humaine, 91057, Evry, France.
17. Quantitative Sciences Unit, Department of Medicine, Stanford University, Stanford, CA, USA.
18. School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy.
19. Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1167-RID-AGE LabEX DISTALZ Risk actors and molecular determinants of ageing diseases, F-59000 Lille, France.
20. Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany.
21. German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany.
22. Cluster of Excellence Cellular Stress Responses in Aging-associated Diseases (CECAD), University of Cologne, Cologne, Germany.
23. Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.
24. Department of Epidemiology, ErasmusMC, Rotterdam, The Netherlands.
25. Nuffield Department of Population Health Oxford University, Oxford, United Kingdom.
26. 1st Department of Neurology, Medical school, Aristotle University of Thessaloniki, Thessaloniki, Makedonia, Greece.
27. Alzheimer's Centre Reina Sofia-CIEN Foundation, Madrid, Spain.
28. Department of Public Health and Caring Sciences / Geriatrics, Uppsala University, Uppsala, Sweden.
29. Krembil Brain Institute, University Health Network, Toronto, Canada.
30. Department of Medicine and Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada.
31. Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy.
32. Institute of Biomedicine, University of Eastern Finland, Joensuu, Kuopio, Eastern Finland, Finland.
33. Division of Psychological Medicine and Clinical Neuroscience, School of Medicine, Cardiff University, Wales, United Kingdom.
34. Department of Neurodegenerative diseases and Geriatric Psychiatry, University Hospital Bonn, Medical Faculty, Bonn, Germany.
35. Department of Psychiatry & Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, San Antonio, TX, USA.
36. NORMENT Centre, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.
37. Institute of Clinical Medicine, University of Oslo, Oslo, Norway.
38. Department of Clinical Sciences and Community Health, University of Milan, 20122 Milan, Italy.
39. Molecular Markers Laboratory, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy.

- 1 40. Centre for Neurodegenerative Disorders, Neurology Unit, Department of Clinical and Experimental Sciences, University of
2 Brescia, Brescia, Italy.
- 3 41. Unit of Neurology, University of Parma and AOU, Parma, Italy.
- 4 42. Department of Neuroscience, Università Cattolica del Sacro Cuore, Rome, Italy.
- 5 43. Neurology Unit, IRCCS Fondazione Policlinico Universitario A. Gemelli, Rome, Italy.
- 6 44. University Bordeaux, Inserm, Bordeaux Population Health Research Center, Bordeaux, France.
- 7 45. Department of Neurology, Bordeaux University Hospital, Bordeaux, France.
- 8 46. Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany.
- 9 47. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
- 10 48. Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.
- 11 49. Inserm, Bordeaux Population Health Research Center, UMR 1219, Univ. Bordeaux, ISPED, CIC 1401-EC, Université de
12 Bordeaux, Bordeaux, France.
- 13 50. CHU de Bordeaux, Pole santé publique, Bordeaux, France.
- 14 51. German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany.
- 15 52. Institute of Cognitive Neurology and Dementia Research (IKND), Otto-von-Guericke University, Magdeburg, Germany.
- 16 53. Neurodegenerative Diseases Unit, Fondazione IRCCS Ca' Granda, Ospedale Policlinico, Milan, Italy.
- 17 54. Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy.
- 18 55. Technical University of Munich, School of Medicine, Klinikum rechts der Isar, Department of Psychiatry and Psychotherapy,
19 Munich, Germany.
- 20 56. Unit for Hereditary Dementias, Theme Aging, Karolinska University Hospital-Solna, 171 64 Stockholm, Sweden.
- 21 57. Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich, University of
22 Zurich, Zurich, Switzerland.
- 23 58. Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland.
- 24 59. Zurich Center for Integrative Human Physiology, University of Zurich, Switzerland.
- 25 60. Université de Paris, EA 4468, APHP, Hôpital Broca, Paris, France.
- 26 61. Department of Geriatric Psychiatry, Central Institute of Mental Health Mannheim, Faculty Mannheim, University of
27 Heidelberg, Germany.
- 28 62. Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany.
- 29 63. Department of Clinical Genetics, VU University Medical Centre, Amsterdam, The Netherlands.
- 30 64. Memory Clinic, Department of Neurology, Charles University, 2nd Faculty of Medicine and Motol University Hospital,
31 Czech Republic.
- 32 65. International Clinical Research Center, St. Anne's University Hospital Brno, Brno, Czech Republic.
- 33 66. Department of Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital of Würzburg,
34 Würzburg, Germany
- 35 67. Department of Radiology&Nuclear medicine, ErasmusMC, Rotterdam, The Netherlands.
- 36 68. Institute of Neurology, Catholic University of the Sacred Heart, Rome, Italy.
- 37 69. Institute of Gerontology and Geriatrics, Department of Medicine and Surgery, University of Perugia, Italy.
- 38 70. Clinic of Neurology, UH "Alexandrovska", Medical University - Sofia, Sofia, Bulgaria.
- 39 71. Faculty of Medicine, University of Lisbon, Portugal.
- 40 72. Institute for Urban Public Health, University Hospital of University Duisburg-Essen, Essen, Germany.
- 41 73. Department of Neuroscience, Psychology, Drug Research and Child Health University of Florence, Florence, Italy.
- 42 74. IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy.
- 43 75. Normandie Univ, UNIROUEN, Inserm U1245 and CHU Rouen, Department of Genetics and CNR-MAJ, F-76000 Rouen,
44 France.
- 45 76. Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm
46 University, Stockholm, Sweden.
- 47 77. Centre for Memory Disturbances, Lab of Clinical Neurochemistry, Section of Neurology, University of Perugia, Perugia,
48 Italy.
- 49 78. Université de Lille, Inserm 1172, CHU Clinical and Research Memory Research Centre (CMRR) of Distalz, Lille, France.
- 50 79. German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany.

- 1 80. Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and
- 2 Berlin Institute of Health, Institute of Psychiatry and Psychotherapy, Hindenburgdamm 30, 12203 Berlin, Germany.
- 3 81. Old Age Psychiatry, Department of Psychiatry, Lausanne University Hospital, Lausanne, Switzerland.
- 4 82. Department of Geriatric Psychiatry, University Hospital of Psychiatry Zürich, Zürich, Switzerland.
- 5 83. Institute for Regenerative Medicine, University of Zürich, Switzerland.
- 6 84. Department of Neuroscience “Rita Levi Montalcini”, University of Torino, Torino, Italy.
- 7 85. Maastricht University, Department of Psychiatry & Neuropsychologie, Alzheimer Center Limburg, Maastricht, The
- 8 Netherlands.
- 9 86. Institute of Social Medicine, Occupational Health and Public Health, University of Leipzig, Leipzig, Germany.
- 10 87. Medical University of Vienna, Department of Psychiatry and Psychotherapy, Vienna, Austria
- 11 88. 1st Department of Neurology, Aiginition Hospital, National and Kapodistrian University of Athens, Medical School, Greece.
- 12 89. Taub Institute for Research in Alzheimer’s Disease and the Aging Brain, The Gertrude H. Sergievsky Center, Department of
- 13 Neurology, Columbia University, New York, New York, United States of America.
- 14 90. LVR-Hospital Essen, Department of Psychiatry and Psychotherapy, Medical Faculty, University of Duisburg-Essen, Essen,
- 15 Germany.
- 16 91. Department for Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Venusberg-Campus 1,
- 17 53127 Bonn, Germany.
- 18 92. Laboratory for Advanced Hematological Diagnostics, Department of Hematology and Stem Cell Transplant, "Vito
- 19 Fazzi"Hospital, Lecce (LE), Italy.
- 20 93. Institute of Clinical Medicine - Neurology, University of Eastern Finland, Finland.
- 21 94. Interdisciplinary Department of Medicine, Geriatric Medicine and Memory Unit, University of Bari “A. Moro, Bari, Italy.
- 22 95. Laboratory of Neuropsychiatry, IRCCS Santa Lucia Foundation, Rome, Italy.
- 23 96. Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, Texas, United States of America.
- 24 97. Department of Biomedical Sciences, University of Cagliari, Italy.
- 25 98. Department of Neurology, ErasmusMC, Rotterdam, The Netherlands.
- 26 99. Neurology, "San Gerardo" hospital, Monza and University of Milano-Bicocca, Italy.
- 27 100. Maastricht University, Department of Psychiatry & Neuropsychologie, Alzheimer Center Limburg, Maastricht, Netherlands.
- 28 101. Department of Psychiatry and Psychotherapy, University Medical Center Goettingen, Goettingen, Germany.
- 29 102. German Center for Neurodegenerative Diseases (DZNE), Goettingen, Germany.
- 30 103. Neurosciences and Signaling Group, Institute of Biomedicine (iBiMED), Department of Medical Sciences, University of
- 31 Aveiro, Aveiro, Portugal.
- 32 104. Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario de Valme, Sevilla, Spain.
- 33 105. Depatamento de Especialidades Quirúrgicas, Bioquímica e Inmunología. Facultad de Medicina. Universidad de Málaga.
- 34 Málaga, Spain
- 35 106. Laboratorio de Genética. Hospital Universitario Central de Asturias, Oviedo, Spain.
- 36 107. Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Spain.
- 37 108. Centro de Biología Molecular Severo Ochoa (UAM-CSIC), Universidad Autónoma de Madrid, Madrid, Spain.
- 38 109. Instituto de Investigacion Sanitaria ‘Hospital la Paz’ (IdIPaz), Madrid, Spain.
- 39 110. Department of Neurology, Il B Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona,
- 40 Barcelona, Spain.
- 41 111. Alzheimer Research Center & Memory Clinic, Andalusian Institute for Neuroscience, Málaga, Spain.
- 42 112. Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología, Instituto de Biomedicina de Sevilla (IBIS),
- 43 Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville, Spain.
- 44 113. Department of Neurology. Hospital Universitario Donostia. San Sebastian, Spain.
- 45 114. Neurosciences Area. Instituto Biodonostia. San Sebastian, Spain.
- 46 115. Fundació Docència i Recerca MútuaTerrassa, Terrassa, Barcelona, Spain.
- 47 116. Memory Disorders Unit, Department of Neurology, Hospital Universitari Mutua de Terrassa, Terrassa, Barcelona, Spain.
- 48 117. Unitat Trastorns Cognitius, Hospital Universitari Santa Maria de Lleida, Lleida, Spain.
- 49 118. Institut de Recerca Biomedica de Lleida (IRBLleida), Lleida, Spain.
- 50 119. Neurological Tissue Bank - Biobanc- Hospital Clinic -IDIBAPS, Barcelona, Spain.
- 51 120. Alzheimer’s disease and other cognitive disorders Unit. Neurology Department, Hospital Clinic , Barcelona, Spain.

- 1 121. Unitat de Genètica Molecular, Institut de Biomedicina de València-CSIC, Valencia, Spain.
- 2 122. Unidad Mixta de Neurología Genética, Instituto de Investigación Sanitaria La Fe, Valencia, Spain.
- 3 123. Neurology Service, Marqués de Valdecilla University Hospital (University of Cantabria and IDIVAL), Santander, Spain.
- 4 124. Departamento de Especialidades Quirúrgicas, Bioquímica e Inmunología. Facultad de Medicina. Universidad de Málaga.
- 5 Málaga, Spain.
- 6 125. Alzheimer's disease and other cognitive disorders unit. Service of Neurology. Hospital Clínic of Barcelona. Institut
- 7 d'Investigacions Biomèdiques August Pi i Sunyer, University of Barcelona, Barcelona, Spain.

1 **Key Points (75-100 word or less)**
2

3 **Question:** Are *APOE* missense variants, other than the common *APOE* alleles $\epsilon 2$ and
4 $\epsilon 4$, associated with AD risk?

5 **Findings:** We meta-analyzed multiple studies including 67,896 Alzheimer's disease (AD)
6 cases, 28,484 proxy-AD cases and 340,306 healthy controls. Two rare missense variants
7 substantially reduced the risk of AD. *APOE*- $\epsilon 3$ [V236E] reported previously but lacking
8 large-scale validation, reduced risk by more than 60%. *APOE*- $\epsilon 4$ [R251G], not previously
9 associated with AD, reduced risk by more than 50% and reached genome-wide
10 significance.

11 **Meaning:** Single amino acid alterations of the *APOE*- $\epsilon 3$ and *APOE*- $\epsilon 4$ isoforms can
12 result in substantial risk reduction for AD. Functional studies examining these variants
13 should elucidate the role of apoE in AD pathogenesis.

14

15

1 **Abstract (350-word limit)**

2 **Importance:** The *APOE*- ϵ 2 and *APOE*- ϵ 4 alleles are, respectively, the strongest
3 protective and risk-increasing genetic variants for late-onset Alzheimer’s disease (AD).
4 However, the mechanisms linking *APOE* to (AD)—particularly the apoE protein’s role in
5 AD pathogenesis and how this is affected by *APOE* variants—remain poorly understood.
6 Identifying missense variants in addition to *APOE*- ϵ 2 and *APOE*- ϵ 4 could provide critical
7 new insights.

8 **Objective:** To determine whether rare missense variants on *APOE* are associated with
9 AD risk.

10 **Design:** Association with case-control status was tested in a sequenced discovery
11 sample (Stage 1) and followed-up in several microarray imputed cohorts as well as the
12 UK Biobank whole-exome sequencing resource using a proxy-AD phenotype (Stages
13 2+3). All data were retrieved between September 2015 and November 2021 and
14 analyzed between April 2021 and November 2021.

15 **Setting:** This study combined case-control, family-based, population-based, and
16 longitudinal AD-related cohorts that recruited referred and volunteer participants.

17 **Participants:** Stage 1 included 37,409 non-unique participants of European or Admixed-
18 European ancestry, with 11,868 cases and 11,934 controls passing analysis inclusion
19 criteria. In Stages 2+3, 475,473 participants were considered across 8 cohorts, of which
20 84,513 cases and proxy-AD cases, and 328,372 controls passed inclusion criteria, and
21 were of European ancestry. Selection criteria were cohort specific, and this study was
22 performed a posteriori on individuals who were genotyped. Among the available
23 genotypes 76,195 were excluded. The number who declined to participate in the original
24 studies was not available.

25 **Main Outcome(s) and Measure(s):** In primary analyses, the AD risk associated with
26 each missense variant was estimated, as appropriate, with either linear-mixed-model
27 regression or logistic regression. In secondary analyses, we estimated associations with
28 age-at-onset using linear-mixed-model regression, and risk of conversion to AD using
29 competing risk regression.

30 **Results:** A total of 544,384 participants (57.4% females, age range 40-110 years old)
31 were analyzed in the primary case-control analysis. Two missense variants were
32 associated with a two to three-fold decreased AD risk: *APOE*- ϵ 4[R251G] (odds ratio,
33 0.44; 95% confidence interval [CI], 0.33-0.59; $P=4.7\times 10^{-8}$) and *APOE*- ϵ 3[V236E] (odds
34 ratio, 0.37; 95% CI, 0.25-0.56; $P=1.9\times 10^{-6}$). Additionally, the cumulative incidence of AD
35 in carriers of these variants was found to grow more slowly with age compared to non-
36 carriers.

37 **Conclusions and Relevance:** We identified a novel variant associated with AD, R251G
38 always co-inherited with ϵ 4, which mitigates the ϵ 4 associated AD risk, and confirmed

1 the protective effect of the V236E variant, always co-inherited with $\epsilon 3$. The location of
2 these variants confirms that the carboxyl-terminal portion of apoE plays an important
3 role in AD pathogenesis. The large risk reductions reported here, suggest that protein
4 chemistry and functional assays of these variants should be pursued as they have the
5 potential to guide drug development targeting *APOE*.

1 **Introduction**

2 Late-onset Alzheimer's disease (AD) is a highly polygenic neurodegenerative disorder
3 with, to date, 75 risk loci associated with AD risk¹. Most of the common single nucleotide
4 polymorphisms (SNPs) at these loci only contribute a small amount to an individual's risk
5 of AD², with the exception of the *APOE*- ϵ 2 and ϵ 4 missense variants that are associated
6 with substantially decreased³ and increased AD risk⁴, respectively. It is estimated that
7 25% of the genetic variance of AD can be attributed to *APOE*- ϵ 2 and *APOE*- ϵ 4⁵. Despite
8 the outsized role of these two common *APOE* alleles, more than 25 years after the initial
9 studies linking them to AD their role in pathogenesis remains ill-defined. Human studies
10 have shown that ϵ 4 speeds, and ϵ 2 slows, the age-related misprocessing of beta-
11 amyloid, though how this occurs at the molecular level remains uncertain^{6,7}. Even the
12 most basic question, does ϵ 4 act via a loss-of-function or gain-of-function mechanism,
13 remains a point of contention⁸. Loss-of-function mutations on *APOE* are exceedingly
14 rare and the sole case report describing a compound heterozygote with two loss-of-
15 function mutations involved a patient who was too young to be informative⁹. The study
16 of additional missense variants on *APOE* may also help to answer this critical question
17 and further elucidate the role of *APOE* in AD. In addition to ϵ 2 and ϵ 4, the only common
18 missense variant (with a minor allele frequency (MAF) > 1%) is Arg145Cys (R145C) an
19 African-ancestry variant always found co-inherited with *APOE*- ϵ 3, which we have shown
20 increases risk for AD¹⁰. The Arg136Ser (R136S) Christchurch variant has recently been
21 posited to play a protective role in early-onset AD related to *PSEN1* mutations, but this
22 study had no statistical genetics support as it was based on data from a single patient¹¹.
23 Finally, strong functional evidence has been marshalled recently to support a protective
24 role for the Val236Glu (V236E) variant, though this was based on data from an earlier
25 case-control study with only ~9,000 subjects^{12,13}, likely underpowered to provide firm
26 estimates of disease risk.

27 On this background, we aimed to investigate, at large scale, the association of rare
28 missense variants on *APOE* with AD risk. We used the Alzheimer's Disease Sequencing
29 Project (ADSP) whole-genome (WGS) and whole-exome sequencing (WES) data as our
30 discovery sample (Stage 1), and sought to replicate significant variants (Stages 2+3) in

1 multiple cohorts using micro-array data imputed on the TOPMed reference panel¹⁴, or
2 by using directly sequenced and genotyped variants from a large Danish general
3 prospective population cohort¹⁵, as well as using the proxy-AD phenotype¹ in the UK
4 Biobank WES data. After filtering, three variants, Leu28Pro (L28P), Val236Glu (V236E),
5 and Arg251Gly (R251G), were tested for their association with AD risk after adjusting for
6 ϵ 2 and ϵ 4 dosages. In complementary analyses, we assessed these associations in an
7 *APOE*-stratified approach to account for the complete linkage disequilibrium of these
8 variants with either the ϵ 2, ϵ 3, or ϵ 4 allele. In secondary analyses, combining Stages 1
9 and 2 datasets, we tested their association with age-at-onset in AD cases and with risk
10 of conversion to AD using competing risk regression.

11

1 **Methods**

2

3 **PARTICIPANTS AND SOURCES OF DATA**

4 Participants or their caregivers provided written informed consent in the original studies.
5 The current study protocol was granted an exemption by the Stanford University
6 institutional review board because the analyses were carried out on deidentified, off-the-
7 shelf data; therefore, additional informed consent was not required. For Stage 1 and
8 Stage 2, phenotypic information and genotypes were obtained from publicly released
9 genome-wide association study datasets assembled by the Alzheimer's Disease
10 Genetics Consortium (ADGC) and derived from WES and WGS data generated by the
11 Alzheimer Disease Sequencing Project (ADSP), with phenotype and genotype
12 ascertainment described elsewhere¹⁶⁻²⁰. The cohorts' queried accession numbers, as
13 well as the sequencing technology or single nucleotide polymorphism (SNP) genotyping
14 platforms are described in **eTables 1 and 2**. Information about Stage 3, which included
15 external replication cohorts and UK Biobank, is provided as a **Supplementary Note**.
16 Briefly, these included EABD-core, EADI, GERAD, DemGene, and GR@ACE/DEGESCO
17 for which phenotype, genotype quality control and imputation have already been
18 described in Bellenguez et al.¹; and CCHS & CGPS *APOE* sequencing and genotyping
19 were described in Rasmussen et al.¹⁵. The following sections describe quality control
20 procedures and ancestry determination applied to the ADSP and ADGC samples
21 respectively used as Stage 1 and Stage 2. The STREGA reporting guidelines were
22 followed.

23

24 **QUALITY CONTROL PROCEDURES**

25 Prior to ancestry, principal components and relatedness determination, in each cohort-
26 platform, variants were excluded based on genotyping rate ($< 95\%$), $MAF < 1\%$, and
27 Hardy-Weinberg equilibrium in controls ($p < 10^{-6}$) using PLINK v1.9²¹. gnomAD²²
28 database-derived information was used to filter out SNPs that met one of the following
29 exclusion criteria^{23,24}: (i) located in a low complexity region, (ii) located within common
30 structural variants ($MAF > 1\%$), (iii) multiallelic SNPs with $MAF > 1\%$ for at least two
31 alternate alleles, (iv) located within a common insertion/deletion, (v) having any flag

1 different than PASS in gnomADv.3, (vi) having potential probe polymorphisms. The latter
2 are defined as SNPs for which the probe may have variable affinity due to the presence
3 of other SNP(s) within 20 bp and with MAF > 1%. Individuals with more than 5%
4 genotype missingness were excluded. Duplicate individuals were identified with KING²⁵
5 and their clinical, diagnostic and pathological data (including age-at-onset of cognitive
6 symptoms, age-at-examination for clinical diagnosis, age-at-last exam, age-at-death),
7 as well as sex, race, and *APOE* genotype were cross-referenced across cohorts.
8 Duplicate entries with irreconcilable phenotype or discordant sex were flagged for
9 exclusion. For individuals with duplicated genotype in sequencing and imputed data, the
10 sequencing entry was used in the Stage 1 discovery set and the imputed entry was not
11 included in the Stage 2 replication set. To apply the *PCAir* and *PCRelate* methods described
12 in the statistical analysis section, we simply considered the intersection of the variants passing
13 quality control in both ADSP WES and ADSP WGS in the discovery, and similarly the intersection
14 of the variants across cohorts genotyping platform in the replication.

15

16

17 **ANCESTRY DETERMINATION**

18 For each cohort, we first determined the ancestry of each individual with SNPWeights
19 v2²⁶ using reference populations from the 1000 Genomes Consortium²⁷. By applying an
20 ancestry percentage cut-off > 75%, the samples were stratified into five super
21 populations: South-Asians, East-Asians, Americans, Africans, and Europeans, and an
22 Admixed group composed of individuals not passing the 75% cut-off in any single
23 ancestry (**eTable 3**)^{10,23}. Since the *APOE* missense variants of interest L28P, V236E, and
24 R251G are too rare to assess reliably in non-European ancestry populations (**eTable 4**),
25 we restricted our analysis to European and Admixed-European individuals. Admixed-
26 European individuals were also included in the main analysis and were part of the
27 Admixed group defined above and had at least 15% European ancestry. We performed
28 sensitivity analyses in increments of 30%, including Admixed-European individuals at
29 45% and 75% cutoffs. The latter corresponding to the super population threshold.

30

1 **IMPUTATION**

2 Each cohort-genotyping platform was imputed on the TOPMed imputation server per
3 ancestry group to obtain an imputation quality (R^2) per ancestry group. We retained
4 cohorts with $R^2 > 0.70$ at rs199768005 for the V236E analyses, and at rs26760661 for
5 the R251G analyses. As there was no significant association signal for rs769452 (L28P)
6 in the Stage 1 primary analysis, we did not check its imputation quality in Stage 2
7 samples.

8

9 **APOE GENOTYPE ASCERTAINMENT**

10 We directed specific attention to the genotyping of the SNPs determining the main *APOE*
11 genotype (rs429358 and rs7412), rs769452-C (*APOE*[L28P]), rs199768005-A
12 (*APOE*[V236E]), and rs267606661-G (*APOE*[R251G]) and follow the procedure described
13 in ¹⁰. Note that Leu28Pro (L28P), Val236Glu (V236E), and Arg251Gly (R251G) are also
14 sometimes respectively referred to as L46P, V254E, and R269G, when the first 18
15 codons of *APOE* encoding a signal peptide are included.

16

17 **SAMPLES ANALYZED**

18 Our discovery sample (Stage 1) was composed of European and Admixed-European
19 ancestry individuals from the ADSP WES and WGS, corresponding to 11,868 AD cases
20 and 11,934 cognitively normal controls (**Table 1**). **eFigure 1** provides a flowchart of the
21 filtering steps leading to the inclusion of these individuals and describes how these
22 datasets were combined. To build a replication sample (Stage 2) for V236E and R251G,
23 we queried for individuals of European and Admixed-European ancestry in all the
24 publicly available microarray genetic datasets that we had access to at the time of the
25 study in July 2021 (**Table 1**). These datasets are largely part of the ADGC and as such
26 this replication will be referred to hereafter as the ADGC replication in Stage 2. After
27 quality control and duplicate removal, 7,768 AD cases and 8,059 controls remained in
28 the ADGC replication sample. **eTable 5** presents the demographics of the remaining AD
29 cases and cognitively unimpaired controls. In Stage 3, we pursued additional replication
30 in external datasets (not publicly available) and in the UK Biobank WES using the proxy-

1 AD phenotype (**Table 1, Supplementary Note**). Overall, the external replications
2 included 36,393 cases and 150,943 controls, and the UK Biobank replication included
3 28,484 proxy-AD cases and 157,436 controls. Across cohorts reported in Table 1, the
4 *APOE* genotype were split as follows: $\epsilon 2/\epsilon 2$: 0.5%, $\epsilon 2/\epsilon 3$: 10.4%, $\epsilon 3/\epsilon 3$: 54.5%, $\epsilon 2/\epsilon 4$:
5 2.5%, $\epsilon 3/\epsilon 4$: 27.7%, $\epsilon 4/\epsilon 4$: 4.4%.

6

7 **STUDY DESIGN & STATISTICAL ANALYSES**

8 In our analysis, we only considered missense variants with a minor allele count above 10
9 in any *APOE* main genotype groups in our next generation sequencing discovery (Stage
10 1) to avoid outlier-confounded effect size estimates²⁸. Three *APOE* missense variants
11 were retained for further analyses: L28P, V236E, and R251G (**eTable 4**). The V236E
12 variant is always co-inherited with *APOE*- $\epsilon 3$, and the L28P and R251G are always co-
13 inherited with *APOE*- $\epsilon 4$ (**eTable 6**). Two variants are co-inherited when they are on the
14 same chromosome copy and close enough to each other that a meiotic crossover event
15 never occurs between them. We thus developed two complementary approaches to take
16 into account these linkage disequilibrium structures. In primary analyses, we estimated
17 the AD risk associated with L28P, V236E, and R251G on case-control diagnoses using
18 linear-mixed-model regression (Stages 1+2, and UK Biobank) and logistic regression
19 (Stage 3) model, adjusted for $\epsilon 2$ and $\epsilon 4$ dosages, in addition to the covariates described
20 below for all analyses. The adjustment by the common $\epsilon 3$ and $\epsilon 4$ *APOE* alleles is
21 necessary because the rare variants tested here are always co-inherited with either the
22 $\epsilon 3$ or $\epsilon 4$ *APOE* allele. In complementary analyses, we also estimated the AD risk
23 associated with V236E and R251G stratified by their associated common *APOE* allele
24 genotype. V236E was assessed in *APOE*- $\epsilon 3/\epsilon 3$ and R251G was assessed in the *APOE*-
25 $\epsilon 3/\epsilon 4$ stratum. An association was considered significant in Stage 1, if it reached a
26 Bonferroni-corrected p-value threshold of 0.017 ($\approx 0.05/3$) in the model adjusted for $\epsilon 2$
27 and $\epsilon 4$ dosages. L28P was not associated with AD risk in this model and was not studied
28 further.

29 Sample sizes and demographics for the stratified analyses are shown in **eTable**
30 **6**. In sensitivity analyses, we estimated AD risk associations for different European

1 ancestry inclusion thresholds. In secondary analyses, combining Stages 1 and 2
2 datasets, we estimated the influence of significant Stage 1 variants on age-at-onset
3 (AAO) in AD cases using linear mixed model regression, and risk of conversion to AD
4 using competing risk regression. In secondary analyses, associations were considered
5 significant when passing the nominal p-value threshold of 0.05. The case-control and
6 age-at-onset analyses used linear mixed model regression available through the
7 *GENESIS* package (v3.12)²⁹. Multivariate competing risk regression and cumulative
8 incidence estimation were implemented using the *cmprsk* package (v2.2)³⁰. In this time-
9 to-event analysis, failure events were defined as age-at-onset for cases (conversion to
10 AD) and age-at-death for controls. Controls without reported death were right censored
11 at age-at-last-visit. Left censoring was set at 50 years old, and younger individuals were
12 excluded from the analysis. All statistical analyses were adjusted for sex and four genetic
13 principal components estimated with the *PC-Air* method³¹ implemented in *GENESIS*.
14 Linear mixed model analyses were additionally covaried by a sparse genetic relationship
15 matrix estimated with the *PC-Relate* method³² implemented in *GENESIS*. Case-control
16 analyses were not adjusted for age given that correcting for age when cases are younger
17 than controls leads to the model incorrectly inferring the age effect on AD risk, resulting
18 in statistical power loss²³.
19 Case-control analyses in Stage 3, external replication cohorts and proxy-AD phenotype
20 in UK Biobank, were implemented to be consistent with the Stage 1 primary analyses.
21 Exact model/analysis details are described in a **Supplementary Note**. For the
22 ADSP/ADGC cohorts, all statistical analyses were performed in R (v4.0.2). All meta-
23 analyses were implemented with a fixed-effect inverse variance weighted design
24 implemented in the *metafor* R package (v.3.0.2)³³.

1 RESULTS

2 In Stage 1 primary analyses, V236E (rs199768005-A) and R251G (rs267606661-
3 G) were associated with a four to five-fold decreased AD risk in non-stratified analyses
4 adjusted for $\epsilon 2$ and $\epsilon 4$ dosages (V236E: OR = 0.23; 95% CI; 0.09-0.56; P = 1.4×10^{-3} ;
5 R251G: OR = 0.20; 95% CI; 0.08-0.49; P = 3.7×10^{-4} , **Figure 1, Table 2**). Similarly, in
6 *APOE*-stratified analyses, V236E was associated with a threefold decreased AD risk in
7 $\epsilon 3/\epsilon 3$ individuals (OR = 0.31; 95% CI; 0.12-0.82; P = 0.02) and R251G was associated
8 with a fivefold decreased AD risk in $\epsilon 3/\epsilon 4$ individuals (OR = 0.17; 95% CI; 0.06-0.48; P
9 = 7.8×10^{-4} , **Table 2**). The L28P variant (rs769452-C) was not associated with AD risk in
10 the non-stratified analyses (odds ratio (OR) = 1.12; 95% confidence interval [CI]; 0.77-
11 1.62; P = 0.56). As such, it was not investigated further.

12 In Stages 2+3, across multiple replication cohorts, the effects of V236E and
13 R251G in non-stratified analyses were concordant and both were significantly
14 associated with AD risk: V236E (OR = 0.42; 95% CI, 0.27-0.66; P= 2.0×10^{-4}) and R251G
15 (OR = 0.48; 95% CI, 0.35-0.66; P= 5.8×10^{-6}). The overall meta-analysis (**Figure 1, Table**
16 **2**) provides robust effect size estimate for these two variants and confirmed their
17 association with a two to three-fold decreased AD risk: V236E (OR = 0.37; 95% CI, 0.25-
18 0.56; P= 1.9×10^{-6}) and R251G (OR = 0.44; 95% CI, 0.33-0.59; P= 4.7×10^{-8}). Similar results
19 were obtained in *APOE*-stratified meta-analyses (**Table 2, eFigure 1**). We further
20 estimated the odds per *APOE* genotype group, using $\epsilon 3/\epsilon 3$ individuals that did not carry
21 V236E as the reference (i.e., odds ratio of *APOE*- $\epsilon 3/\epsilon 3$ individuals equals 1), by meta-
22 analyzing the ADSP discovery and ADGC replication cohorts. Compared to the reference
23 $\epsilon 3/\epsilon 3$ group, $\epsilon 3/\epsilon 3$ [V236E] and $\epsilon 3/\epsilon 4$ [R251G] individuals had AD risk lower than or similar
24 to $\epsilon 2/\epsilon 3$ (**Figure 2**).

25 Results of sensitivity analyses evaluating different European ancestry cutoffs are
26 shown in (**eTable 8, eFigure 2**). Briefly, the results remained unchanged when selecting
27 admixed ancestry individuals with at least 45% European ancestry, or when restricting
28 the analysis to European ancestry individuals (75% cutoff). We note that the odds ratio
29 in the combined ADSP/ADGC datasets for V236E and R251G remain unchanged at

1 different ancestry cutoffs. For example, using an ancestry cutoff at 75%, the non-
2 stratified meta-analysis yielded an odds ratio of 0.27 (95% CI, 0.12 to 0.58; $P = 8.6 \times 10^{-4}$)
3 ⁴) for V236E as compared to an odds ratio of 0.26 using a cutoff of 15%. Similar
4 observations were made for the R251G variant. As additional supplementary analyses,
5 we assessed the effect of the inclusion of “all dementia” (rather than AD specifically) in
6 the CCHS & CGPS dataset and we estimated the significance without including UK
7 Biobank. Overall, the significance of the results slightly improved when including a
8 broader dementia category (e.g. R251G, OR= 0.44; 95% CI, 0.33-0.59; $P=3.5 \times 10^{-8}$,
9 **eTable 9**). While removing UK Biobank proxy-AD phenotype samples reduced the
10 significance of our results slightly, the ORs became slightly more protective (e.g. R251G,
11 OR= 0.39; 95% CI, 0.27-0.56; $P=1.2 \times 10^{-7}$, **eTable 10**).

12 In secondary analyses, including data from Stages 1+2, we considered the meta-
13 analysis of ADSP/ADGC samples (**eTable 5**). In non-*APOE* stratified analyses adjusted
14 for $\epsilon 2$ and $\epsilon 4$ dosages (**eTable 7**), V236E carriers had an age-at-AD-onset on average
15 10.5 years older than non-carriers ($\beta = 10.64$; 95% CI, 1.78 to 19.49; $P = 0.02$) and slower
16 incidence with age (HR = 0.30; 95% CI; 0.12-0.76; $P = 0.01$). While R251G’s association
17 with age-at-onset was not significant ($\beta = 0.97$; 95% CI, -2.96 to 4.91; $P = 0.63$) and its
18 association with reduced AD incidence with age was just nominally significant (HR =
19 0.67; 95% CI; 0.46-0.97; $P = 0.04$). In *APOE*-stratified analyses (**eTable 7**), a similar effect
20 of V236E on age-at-AD-onset was observed in $\epsilon 3/\epsilon 3$ ($\beta = 10.93$; 95% CI, 1.06 to 20.81;
21 $P = 0.03$). R251G carriers had an age-at-AD-onset on average 6 years older than non-
22 carriers in $\epsilon 3/\epsilon 4$ but this association was only trending towards significance ($\beta = 6.04$;
23 95% CI, -0.71 to 12.79; $P = 0.08$). The competing risk results emphasized that the
24 cumulative incidence of AD in $\epsilon 3/\epsilon 3$ participants grows slower with age in individuals
25 carrying the V236E variant (hazard ratio [HR] = 0.40; 95% CI; 0.17-0.97; $P = 0.04$), and
26 similarly in $\epsilon 3/\epsilon 4$ participants carrying the R251G variant (HR = 0.26; 95% CI; 0.13-0.54;
27 $P = 2.9 \times 10^{-4}$).

28

1 DISCUSSION

2 We have shown that two missense variants V236E and R251G are each
3 associated with a more than 2-fold reduction in AD risk (**Figure 2**). These variants have
4 an allele frequency of less than 0.1% in gnomAD v3.1, even when restricting this
5 frequency estimate to Europeans (**eTable 4**). Due to their rarity and linkage disequilibrium
6 with the common *APOE* - ϵ 3 and - ϵ 4 alleles, they have not been identified in prior
7 genome-wide association studies¹. The protective effect of V236E has already been
8 reported in a smaller prior study focused on *APOE*¹³ and was suggestive in a population-
9 based study¹⁵, but we validated this finding here in a large-scale genomic study and
10 provide an improved estimate of its effect size. The association of R251G with AD risk
11 has not been previously reported. This variant, carried on the same haplotype as ϵ 4, is
12 the first *APOE* variant found to mitigate the AD risk attributable to the ϵ 4 isoform of the
13 apoE protein. Notably, having R251G in association with *APOE*- ϵ 4 results in a risk
14 estimate similar to *APOE*- ϵ 2, as shown in **Figure 2** where *APOE*- ϵ 3/ ϵ 4[R251G] and
15 *APOE*- ϵ 2/ ϵ 3 have an equivalent odds ratio. Our study has several limitations (i) the V236E
16 association was not genome-wide significant, (ii) we included the UKB dataset that does
17 not include a direct clinical diagnosis of AD, (iii) due to the paucity of variant carriers in
18 non-European ancestries we did not assess these variants in other ancestries (although
19 they can be found in African-Americans and Admixed-Latinos based on gnomAD
20 estimates (**eTable 4**)). These three caveats point to the need for further confirmation of
21 these variants as available AD datasets grow and become more ancestrally diverse.

22 Regarding potential mechanisms driving these associations, it is notable that
23 these two variants are on apoE's C-terminal domain. The common *APOE*- ϵ 2 and *APOE*-
24 ϵ 4 alleles are located on the N-terminal domain of the protein near the receptor-binding
25 region. Their outsized role in AD risk has, understandably, focused attention on the N-
26 terminal domain and the differential capacity of these alleles to, for example, bind apoE's
27 receptors^{34,35}. The current results add support to studies suggesting that the C-terminal
28 domain is also of critical importance for AD pathogenesis³⁶⁻³⁸. R251G is located within
29 apoE's lipid-binding region (amino acid residues 244 to 272), while V236E is adjacent to

1 this region⁸. A recent publication provided evidence for the protectiveness of V236E
2 against AD pathology and explored the functional mechanism supporting its protective
3 role¹². The lipid-binding region, with its abundance of non-polar residues, is thought to
4 be a region that can foster oligomerization³⁹⁻⁴¹. Switching a non-polar valine for an acidic
5 glutamic acid might be predicted to reduce the hydrophobicity of this region and reduce
6 its tendency to oligomerize. Notably, the authors showed reduced levels of insoluble A β
7 and apoE aggregates in the brain of V236E carriers compared to non-carriers¹². In 5xFAD
8 mice, they observed that *APOE- ϵ 3[V236E]* reduced A β deposition, plaque-associated
9 immune response, and neuritic dystrophy around amyloid plaques¹². Chemically, they
10 noted that *APOE- ϵ 3[V236E]* primarily remains as a monomer and is less likely to form
11 oligomers compared to the canonical *APOE- ϵ 3* allele¹². This propensity of V236E to
12 reduce apoE aggregation was also observed when this variant was introduced on an
13 *APOE- ϵ 4* allele. It is worth noting, however, that V236E also appears to increase
14 dimerization (see their Figure S10¹²), which may impact apoE's ability to bind to its
15 receptors⁴²⁻⁴⁴.

16 Given that R251G is located squarely in the lipid-binding region of the protein, it
17 is possible that R251G confers a protective effect by reducing apoE's ability to form
18 insoluble oligomers. The switch from a charged arginine amino acid to a non-polar
19 glycine might, however, be expected to increase rather than decrease oligomerization.
20 Changes in this region could also enhance apoE- ϵ 4's ability to bind lipids rendering it
21 more like ϵ 3 or ϵ 2 in this capacity⁴⁵. Alternatively, the introduction of glycine could disrupt
22 the alpha-helix structure of the C-terminal impacting apoE- ϵ 4's hypothesized N-
23 terminal/C-terminal domain interaction^{34,35}. In any case, pending protein chemistry
24 experiments exploring potential structural and functional changes, the mechanism
25 underlying the substantial protective effect of R251G remains to be elucidated.

26 Our work, performed on the largest available sample to-date for these two
27 variants, has validated the protective effect of the V236E variant and has uncovered a
28 novel protective missense variant on *APOE- ϵ 4*. Each variant has a substantial effect on
29 reducing the risk of AD. While some compelling functional data suggest that V236E
30 confers protection by reducing oligomerization of apoE, there are alternative

1 mechanisms that merit consideration (increasing dimerization, for one). The protective
2 mechanism of R251G remains unexplored but finding a single amino acid substitution
3 that renders the *APOE-ε4* allele protective supports the idea that *APOE-ε4*-specific
4 treatments are worth exploring^{46,47}. We anticipate that the findings reported here will
5 spark additional mechanistic work on apoE's role in AD pathogenesis.
6

1 **FUNDING AND ACKNOWLEDGMENTS**

2
3 This work was supported by the National Institute of Health and National Institute of
4 Aging grants AG060747 (MDG), AG066206 (ZH), AG066515 (ZH, MDG), the European
5 Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-
6 Curie (grant agreement No. 890650, YLG), the Alzheimer’s Association (AARF-20-
7 683984, MEB), and the Iqbal Farrukh and Asad Jamal Fund, a grant from the EU Joint
8 Programme – Neurodegenerative Disease Research (European Alzheimer DNA BioBank,
9 EADB; JPND). Inserm UMR1167 is also funded by the Inserm, Institut Pasteur de Lille,
10 Lille Métropole Communauté Urbaine, and the French government’s LABEX DISTALZ
11 program (development of innovative strategies for a transdisciplinary approach to
12 Alzheimer’s disease). EADB thank the study participants, researchers, and staff for
13 collecting and contributing to the data, the high-performance computing service at the
14 University of Lille, and the staff at CEA-CNRGH for their help with sample preparation
15 and genotyping, and technical assistance. Additional funders of individual investigators
16 and institutions who contributed to data collection and genotyping are provided in the
17 **Supplemental Online Content.**

References

1. Bellenguez C, Küçükali F, Jansen I, et al. New insights on the genetic etiology of Alzheimer's and related dementia. *medRxiv*. Published online January 1, 2020:2020.10.01.20200659. doi:10.1101/2020.10.01.20200659
2. de Rojas I, Moreno-Grau S, Tesi N, et al. Common variants in Alzheimer's disease and risk stratification by polygenic risk scores. *Nat Commun*. 2021;12(1):3417. doi:10.1038/s41467-021-22491-8
3. Corder EH, Saunders AM, Risch NJ, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nature Genetics*. 1994;7(2):180-184. doi:10.1038/ng0694-180
4. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993;261(5123):921-923. doi:10.1126/science.8346443
5. 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. doi:10.1016/j.neurobiolaging.2016.02.024
6. Morris JC, Roe CM, Xiong C, et al. APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Annals of Neurology*. 2010;67(1):122-131. doi:10.1002/ana.21843
7. Castellano JM, Kim J, Stewart FR, et al. Human apoE isoforms differentially regulate brain amyloid- β peptide clearance. *Sci Transl Med*. 2011;3(89):89ra57. doi:10.1126/scitranslmed.3002156
8. Belloy ME, Napolioni V, Greicius MD. A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. *Neuron*. 2019;101(5):820-838. doi:10.1016/j.neuron.2019.01.056
9. Mak ACY, Pullinger CR, Tang LF, et al. Effects of the absence of apolipoprotein e on lipoproteins, neurocognitive function, and retinal function. *JAMA Neurol*. 2014;71(10):1228-1236. doi:10.1001/jamaneurol.2014.2011
10. Le Guen Y, Belloy ME, Eger SJ, et al. *APOE Missense Variant R145C Is Associated with Increased Alzheimer's Disease Risk in African Ancestry Individuals with the APOE E3/E4 Genotype.*; 2021:2021.10.20.21265141. doi:10.1101/2021.10.20.21265141
11. Arboleda-Velasquez JF, Lopera F, O'Hare M, et al. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. *Nat Med*. 2019;25(11):1680-1683. doi:10.1038/s41591-019-0611-3

- 1 12. Liu CC, Murray ME, Li X, et al. APOE3-Jacksonville (V236E) variant reduces self-
2 aggregation and risk of dementia. *Science Translational Medicine*.
3 2021;13(613):eabc9375. doi:10.1126/scitranslmed.abc9375
- 4 13. Medway CW, Abdul-Hay S, Mims T, et al. ApoE variant p.V236E is associated
5 with markedly reduced risk of Alzheimer's disease. *Molecular Neurodegeneration*.
6 2014;9(1):11. doi:10.1186/1750-1326-9-11
- 7 14. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes
8 from the NHLBI TOPMed Program. *Nature*. 2021;590(7845):290-299.
9 doi:10.1038/s41586-021-03205-y
- 10 15. Rasmussen KL, Tybjerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. APOE
11 and dementia – resequencing and genotyping in 105,597 individuals. *Alzheimer's &*
12 *Dementia*. 2020;16(12):1624-1637. doi:10.1002/alz.12165
- 13 16. Beecham GW, Bis JC, Martin ER, et al. The Alzheimer's Disease Sequencing
14 Project: Study design and sample selection. *Neurol Genet*. 2017;3(5):e194.
15 doi:10.1212/NXG.000000000000194
- 16 17. Weiner MW, Aisen PS, Jack CR, et al. The Alzheimer's Disease Neuroimaging
17 Initiative: Progress report and future plans. *Alzheimer's & Dementia*. 2010;6(3):202-
18 211.e7. doi:10.1016/j.jalz.2010.03.007
- 19 18. Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS.
20 Overview and findings from the rush Memory and Aging Project. *Current Alzheimer*
21 *research*. 2012;9(6):646-663. doi:10.2174/156720512801322663
- 22 19. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed
23 Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and
24 lipid processing. *Nature Genetics*. 2019;51(3):414-430. doi:10.1038/s41588-019-
25 0358-2
- 26 20. Kunkle BW, Schmidt M, Klein HU, et al. Novel Alzheimer Disease Risk Loci and
27 Pathways in African American Individuals Using the African Genome Resources
28 Panel: A Meta-analysis. *JAMA Neurol*. 2021;78(1):102.
29 doi:10.1001/jamaneurol.2020.3536
- 30 21. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-
31 generation PLINK: rising to the challenge of larger and richer datasets.
32 *GigaScience*. 2015;4(1):7. doi:10.1186/s13742-015-0047-8
- 33 22. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum
34 quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
35 doi:10.1038/s41586-020-2308-7

- 1 23. Le Guen Y, Belloy ME, Napolioni V, et al. A novel age-informed approach for
2 genetic association analysis in Alzheimer's disease. *Alzheimer's Research &*
3 *Therapy*. 2021;13(1):72. doi:10.1186/s13195-021-00808-5
- 4 24. Le Guen Y, Napolioni V, Belloy ME, et al. Common X-Chromosome Variants Are
5 Associated with Parkinson Disease Risk. *Annals of Neurology*. 2021;90(1):22-34.
6 doi:10.1002/ana.26051
- 7 25. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust
8 relationship inference in genome-wide association studies. *Bioinformatics*.
9 2010;26(22):2867-2873. doi:10.1093/bioinformatics/btq559
- 10 26. Chen CY, Pollack S, Hunter DJ, Hirschhorn JN, Kraft P, Price AL. Improved
11 ancestry inference using weights from external reference panels. *Bioinformatics*.
12 2013;29(11):1399-1406. doi:10.1093/bioinformatics/btt144
- 13 27. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic
14 variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393
- 15 28. Bis JC, Jian X, Kunkle BW, et al. Whole exome sequencing study identifies novel
16 rare and common Alzheimer's-Associated variants involved in immune response
17 and transcriptional regulation. *Molecular Psychiatry*. 2020;25(8):1859-1875.
18 doi:10.1038/s41380-018-0112-7
- 19 29. Gogarten SM, Sofer T, Chen H, et al. Genetic association testing using the
20 GENESIS R/Bioconductor package. *Bioinformatics*. 2019;35(24):5346-5348.
21 doi:10.1093/bioinformatics/btz567
- 22 30. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a
23 Competing Risk. *Journal of the American Statistical Association*. 1999;94(446):496-
24 509. doi:10.1080/01621459.1999.10474144
- 25 31. Conomos MP, Miller MB, Thornton TA. Robust Inference of Population Structure
26 for Ancestry Prediction and Correction of Stratification in the Presence of
27 Relatedness. *Genetic Epidemiology*. 2015;39(4):276-293.
28 doi:https://doi.org/10.1002/gepi.21896
- 29 32. Conomos MP, Laurie CA, Stilp AM, et al. Genetic Diversity and Association
30 Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community
31 Health Study/Study of Latinos. *The American Journal of Human Genetics*.
32 2016;98(1):165-184. doi:10.1016/j.ajhg.2015.12.001
- 33 33. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package.
34 *Journal of Statistical Software*. 2010;36(1):1-48. doi:10.18637/jss.v036.i03

- 1 34. Huang Y, Weisgraber KH, Mucke L, Mahley RW. Apolipoprotein E: diversity of
2 cellular origins, structural and biophysical properties, and effects in Alzheimer's
3 disease. *J Mol Neurosci*. 2004;23(3):189-204. doi:10.1385/JMN:23:3:189
- 4 35. Huang Y, Mahley RW. Apolipoprotein E: structure and function in lipid
5 metabolism, neurobiology, and Alzheimer's diseases. *Neurobiol Dis*. 2014;72 Pt
6 A:3-12. doi:10.1016/j.nbd.2014.08.025
- 7 36. Harris FM, Brecht WJ, Xu Q, et al. Carboxyl-terminal-truncated apolipoprotein
8 E4 causes Alzheimer's disease-like neurodegeneration and behavioral deficits in
9 transgenic mice. *PNAS*. 2003;100(19):10966-10971. doi:10.1073/pnas.1434398100
- 10 37. Bien-Ly N, Andrews-Zwilling Y, Xu Q, Bernardo A, Wang C, Huang Y. C-
11 terminal-truncated apolipoprotein (apo) E4 inefficiently clears amyloid- β (A β) and
12 acts in concert with A β to elicit neuronal and behavioral deficits in mice. *PNAS*.
13 2011;108(10):4236-4241. doi:10.1073/pnas.1018381108
- 14 38. Huang YWA, Zhou B, Wernig M, Südhof TC. ApoE2, ApoE3, and ApoE4
15 Differentially Stimulate APP Transcription and A β Secretion. *Cell*. 2017;168(3):427-
16 441.e21. doi:10.1016/j.cell.2016.12.044
- 17 39. Choy N, Raussens V, Narayanaswami V. Inter-molecular coiled-coil formation in
18 human apolipoprotein E C-terminal domain. *J Mol Biol*. 2003;334(3):527-539.
19 doi:10.1016/j.jmb.2003.09.059
- 20 40. Westerlund JA, Weisgraber KH. Discrete carboxyl-terminal segments of
21 apolipoprotein E mediate lipoprotein association and protein oligomerization. *J Biol*
22 *Chem*. 1993;268(21):15745-15750.
- 23 41. Flowers SA, Rebeck GW. APOE in the normal brain. *Neurobiol Dis*.
24 2020;136:104724. doi:10.1016/j.nbd.2019.104724
- 25 42. Dyer CA, Cistola DP, Parry GC, Curtiss LK. Structural features of synthetic
26 peptides of apolipoprotein E that bind the LDL receptor. *J Lipid Res*. 1995;36(1):80-
27 88.
- 28 43. Weisgraber KH, Shinto LH. Identification of the disulfide-linked homodimer of
29 apolipoprotein E3 in plasma. Impact on receptor binding activity. *J Biol Chem*.
30 1991;266(18):12029-12034.
- 31 44. Minami SS, Cordova A, Cirrito JR, et al. ApoE mimetic peptide decreases A β
32 production in vitro and in vivo. *Mol Neurodegener*. 2010;5:16. doi:10.1186/1750-
33 1326-5-16

- 1 45. Minagawa H, Gong JS, Jung CG, et al. Mechanism Underlying Apolipoprotein E
2 (ApoE) Isoform-dependent Lipid Efflux From Neural Cells in Culture. *J Neurosci*
3 *Res.* 2009;87(11):2498-2508. doi:10.1002/jnr.22073
- 4 46. Zhao N, Liu CC, Qiao W, Bu G. Apolipoprotein E, Receptors, and Modulation of
5 Alzheimer's Disease. *Biol Psychiatry.* 2018;83(4):347-357.
6 doi:10.1016/j.biopsych.2017.03.003
- 7 47. Williams T, Borchelt DR, Chakrabarty P. Therapeutic approaches targeting
8 Apolipoprotein E function in Alzheimer's disease. *Mol Neurodegener.* 2020;15(1):8.
9 doi:10.1186/s13024-020-0358-9

10

11

1 **Data availability**

2 Data used in preparation of this manuscript can be obtained upon application at:

- 3 - dbGaP (https://www.ncbi.nlm.nih.gov/gap/advanced_search/)
- 4 - NIAGADS and NIAGADS DSS (<https://www.niagads.org/>)
- 5 - LONI (<https://ida.loni.usc.edu/>)
- 6 - Synapse (<https://adknowledgeportal.synapse.org/>)
- 7 - RADc Rush (<https://www.radc.rush.edu/>)
- 8 - NACC (<https://naccdata.org/>)
- 9 - UK Biobank (<https://biobank.ndph.ox.ac.uk/showcase/>)

10 **eTables 1 and 2** provide the details of repositories and accession number per cohort-
11 platform group. UK Biobank WES data were analyzed under Application Number 45420.

12

Table 1. Demographics per *APOE* genotype. DX: diagnosis, CN: cognitively normal, AD: Alzheimer's disease, N: number of individuals, %Females: percentage of female individuals, μ and σ : mean age and standard deviation. UK Biobank demographics are not reported in this table since cases correspond to proxy-AD phenotype mostly relying on self-report of first-degree relatives' diagnosis without age-at-onset being specified.

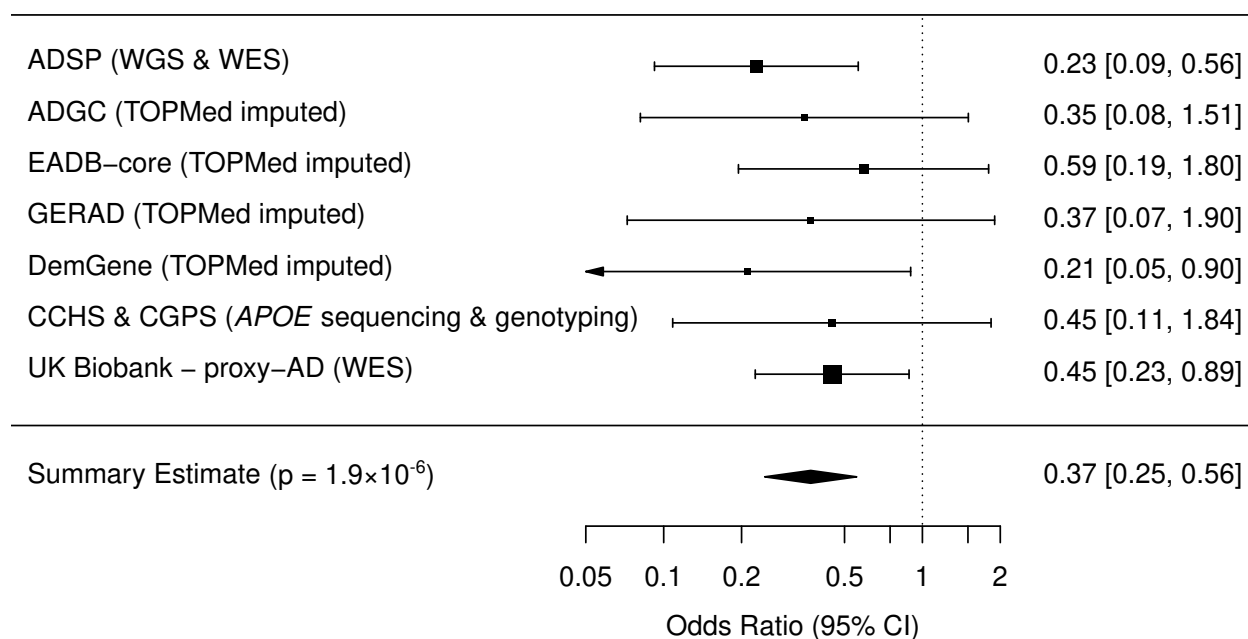
Sample DX	N	APOE ε2/ε2		APOE ε2/ε3		APOE ε3/ε3		APOE ε2/ε4		APOE ε3/ε4		APOE ε4/ε4	
		N (%Females)	Age μ(σ)	N (%Females)	Age μ(σ)	N (%Females)	Age μ(σ)	N (%Females)	Age μ(σ)	N (%Females)	Age μ(σ)	N (%Females)	Age μ(σ)
ADSP	CN 11,934	73(54.8%)	82.6(8.3)	1481(62.4%)	83.0(8.0)	7429(62.4%)	82.3(8.1)	195(70.3%)	79.8(8.9)	2561(62.1%)	79.7(8.2)	195(63.1%)	76.6(7.5)
	AD 11,868	29(58.6%)	82.5(6.9)	583(63.3%)	80.1(9.7)	5313(60.9%)	77.0(10.1)	258(61.2%)	75.3(8.2)	4919(58.0%)	73.2(8.5)	766(53.0%)	67.9(8.1)
ADGC	CN 8,059	56(46.4%)	79.1(10.2)	978(64.3%)	76.2(9.5)	4795(61.9%)	74.5(9.4)	209(63.2%)	73.8(10.1)	1847(61.9%)	71.4(10.1)	174(60.9%)	68.7(9.3)
	AD 7,768	10(60.0%)	72.5(8.2)	323(56.0%)	75.8(10.4)	2494(63.6%)	74.7(10.5)	237(63.3%)	75.7(8.8)	3258(63.2%)	73.0(8.6)	1446(57.4%)	69.7(7.2)
EADB-core	CN 21,160	121(59.5%)	68.6(13.2)	2503(58.2%)	66.8(15.1)	13365(57.8%)	67.0(14.5)	396(55.6%)	66.7(13.3)	4390(55.7%)	66.3(13.6)	385(55.1%)	64.2(12.6)
	AD 19,873	27(51.9%)	76.4(11.7)	877(59.5%)	74.2(11.2)	8285(61.9%)	72.9(11.0)	435(66.0%)	73.2(10.7)	8003(63.0%)	71.7(9.7)	2246(57.4%)	67.6(8.8)
GR@ACE	CN 8,539	33(57.6%)	53.1(17.6)	858(52.2%)	57.5(18.7)	6005(50.1%)	56.7(18.0)	99(49.5%)	56.7(17.6)	1459(49.8%)	56.7(17.6)	85(43.5%)	54.9(14.8)
	AD 7,355	16(84.6%)	84.6(3.5)	389(70.4%)	81.4(8.1)	3840(70.4%)	80.9(7.9)	115(73.0%)	78.7(7.4)	2590(69.8%)	78.7(7.4)	405(64.7%)	74.8(7.3)
EADI	CN 6,331	38(52.6%)	82.6(7.5)	772(59.2%)	81.0(7.5)	4247(60.8%)	80.1(7.7)	109(60.6%)	78.8(7.1)	1106(59.2%)	79.0(7.6)	59(71.2%)	77.1(6.7)
	AD 2,397	7(85.7%)	79.3(6.0)	128(68.8%)	78.0(10.8)	1078(65.3%)	76.5(10.6)	71(59.2%)	73.4(8.8)	888(66.0%)	72.6(9.2)	225(64.9%)	68.1(7.0)
GERAD	CN 7,007	47(55.3%)	49.3(11.0)	853(50.1%)	51.5(12.6)	4127(51.9%)	50.9(11.9)	180(51.7%)	49.8(10.9)	1627(51.8%)	49.9(10.9)	173(49.7%)	49.9(11.0)
	AD 2,989	10(60.0%)	81.2(9.7)	140(62.9%)	79.3(11.3)	1092(62.0%)	79.3(9.6)	90(63.3%)	80.4(7.6)	1306(64.2%)	77.7(8.9)	351(62.4%)	74.2(8.4)
DemGene	CN 5,911	32(34.4%)	68.7(11.2)	685(49.1%)	69.2(12.4)	3236(47.6%)	68.9(11.0)	167(45.5%)	70.6(10.6)	1595(48.2%)	67.3(10.5)	196(44.4%)	64.7(11.0)
	AD 1,687	5(40.0%)	74.0(1.4)	72(58.3%)	71.6(10.6)	537(66.9%)	73.7(9.6)	43(72.1%)	75.4(7.0)	769(66.6%)	72.2(8.4)	261(61.7%)	69.3(8.1)
CCHS & CGPS	CN 101,995	705(54.9%)	57.0(13.2)	12818(55.1%)	57.6(13.6)	57115(54.8%)	57.5(13.4)	2936(55.4%)	56.8(13.0)	25616(54.9%)	56.7(12.8)	2778(57.6%)	55.3(12.7)
	AD 2,092	12(50.0%)	72.6(5.3)	129(53.5%)	73.3(8.4)	844(58.8%)	73.3(8.4)	70(61.4%)	71.2(8.0)	821(62.4%)	70.9(8.0)	216(56.9%)	68.8(7.9)

Table 2. V236E and R251G are associated with a decreased AD risk. The significance of their association with AD risk is equivalent in non-stratified analyses adjusted by *APOE* ϵ 2 and ϵ 4 dosages, and in *APOE*-stratified analysis considering the main *APOE* genotype group with the most carriers for each variant, namely ϵ 3/ ϵ 3 and ϵ 3/ ϵ 4 respectively for V236E and R251G.

N: Number of individuals, MAC: Minor allele count, OR: odds ratio, 95% CI: 95% confidence interval, P: p-value.

	Sample	AD Case-Control Regression Non-stratified				AD Case-Control Regression APOE-Stratified			
		N	MAC	OR [95% CI]	P	N	MAC	OR [95% CI]	P
V236E (all APOE (left) and ε3/ε3 only (right))	<i>ADSP</i>	23,427	20	0.23 [0.09; 0.56]	1.4E-03	12,604	17	0.31 [0.12; 0.82]	0.020
	<i>ADGC imputed</i>	11,652	9	0.35 [0.08; 1.51]	0.16	5,741	10	0.40 [0.1; 1.57]	0.19
	<i>EADB-core</i>	41,033	27.17	0.59 [0.19; 1.80]	0.34	21,650	21.28	0.53 [0.15; 1.92]	0.30
	<i>GERAD</i>	9,996	17.72	0.37 [0.07; 1.90]	0.18	5,219	9.43	0.77 [0.10; 6.06]	0.78
	<i>DemGene</i>	7,598	58.68	0.21 [0.05; 0.90]	8.5E-03	3,773	35.88	0.56 [0.13; 2.46]	0.40
	<i>CCHS & CGPS</i>	104,084	240	0.45 [0.11; 1.84]	0.23	57,955	191	0.18 [0.01; 2.97]	0.27
	<i>UKB proxy-AD</i>	185,741	277	0.45 [0.23; 0.89]	0.021	109,120	219	0.47 [0.21; 1.04]	0.063
	Meta-analysis	383,531	649.57	0.37 [0.25; 0.56]	1.9E-06	216,062	503.59	0.43 [0.27; 0.69]	4.4E-04
R251G (all APOE (left) and ε3/ε4 only (right))	<i>ADSP</i>	23,314	26	0.20 [0.08; 0.49]	3.7E-04	7,335	18	0.17 [0.06; 0.48]	7.8E-04
	<i>ADGC imputed</i>	14,134	29	0.29 [0.12; 0.70]	5.8E-03	4,630	16	0.19 [0.07; 0.54]	1.7E-03
	<i>EADB-core</i>	41,033	59.16	0.51 [0.26; 0.99]	0.049	12,393	40.27	0.34 [0.15; 0.76]	7.8E-03
	<i>GR@ACE</i>	15,894	21.27	0.35 [0.12; 1.01]	0.049	4,049	17.81	0.22 [0.06; 0.77]	0.011
	<i>EADI</i>	8,728	19.21	0.68 [0.22; 2.09]	0.49	1,994	13.32	1.14 [0.32; 4.04]	0.84
	<i>GERAD</i>	9,996	23.17	0.50 [0.17; 1.47]	0.18	2,933	16.82	0.57 [0.18; 1.88]	0.34
	<i>CCHS & CGPS</i>	104,087	105	0.41 [0.10; 2.72]	0.23	26,437	75	0.33 [0.05; 2.43]	0.28
	<i>UKB proxy-AD</i>	185,735	335	0.57 [0.34; 0.98]	0.041	43,820	262	0.67 [0.36; 1.22]	0.19
Meta-analysis	402,921	617.81	0.44 [0.33; 0.59]	4.7E-08	103,591	459.22	0.41 [0.29; 0.57]	3.2E-07	

APOE V236E, rs199768005



APOE R251, rs267606661

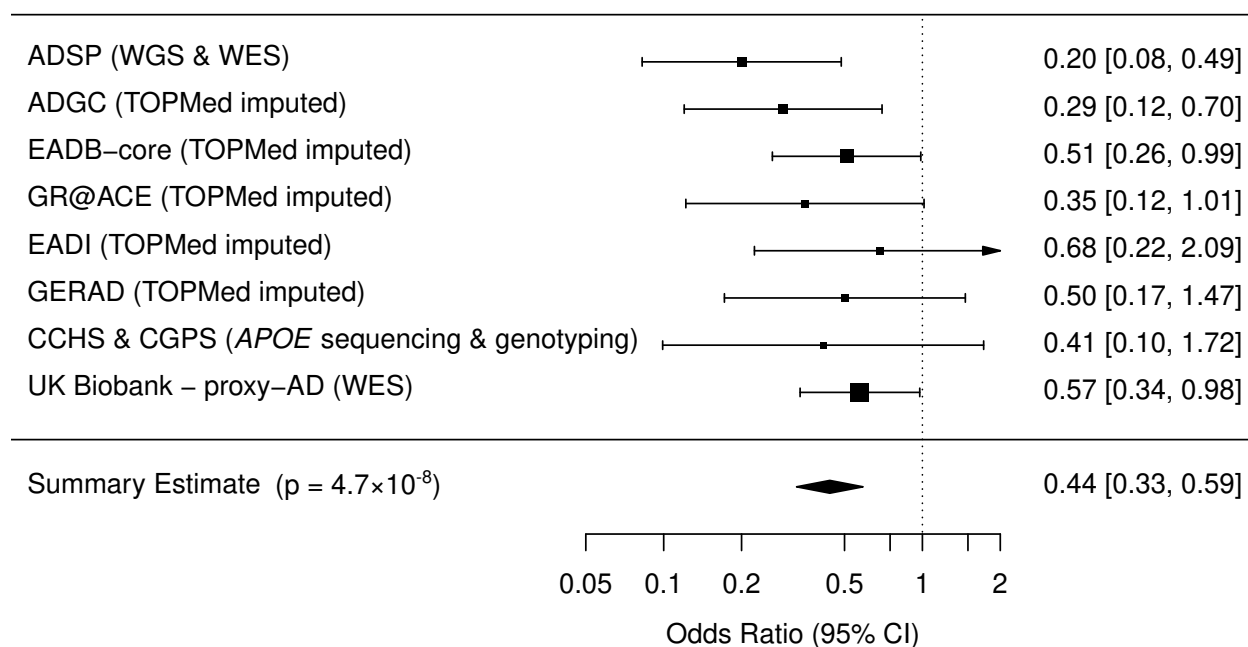


Figure 1. V236E and R251G are associated with decreased AD risk across all cohorts. Forest plots show the results for the non-*APOE* stratified analyses adjusted by $\epsilon 2$ and $\epsilon 4$ dosages. **eFigure 1** presents equivalent forest plots for these two variants in the *APOE*-stratified sensitivity analyses, showing consistent findings.

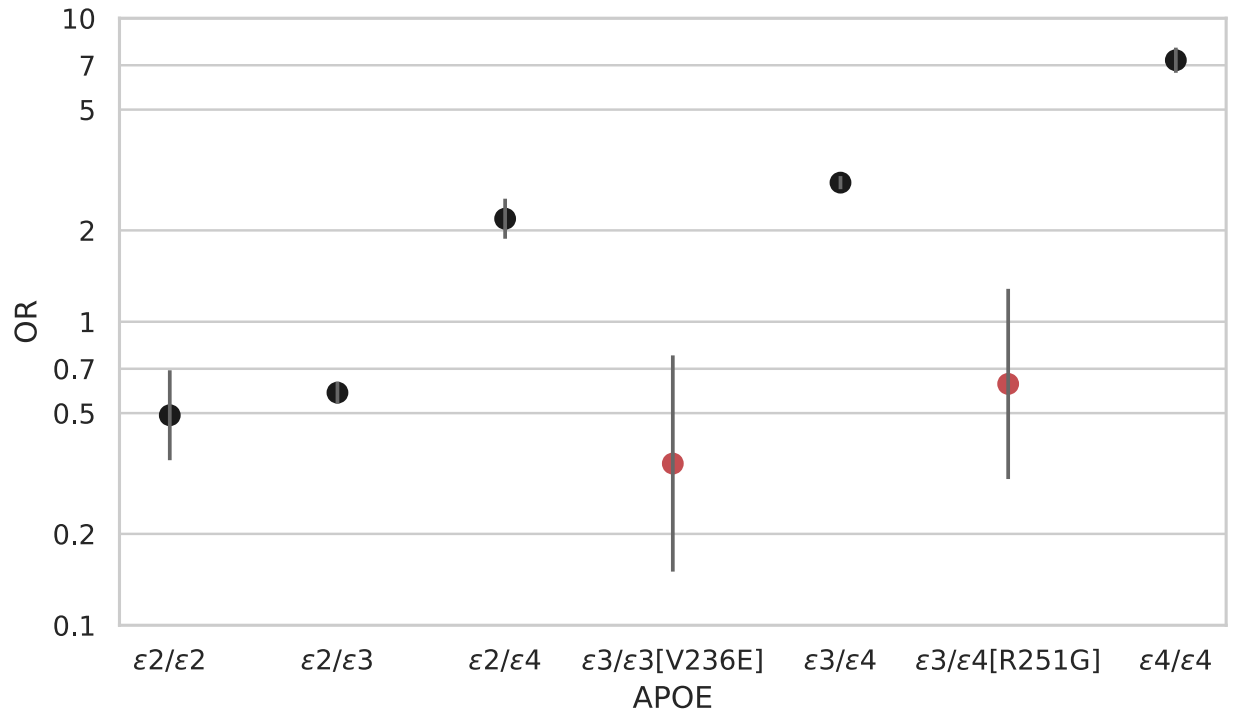


Figure 2. *APOE* $\epsilon 3/\epsilon 3[V236E]$ and *APOE* $\epsilon 3/\epsilon 4[R251G]$ have a risk equivalent to $\epsilon 2/\epsilon 3$ carriers. Alzheimer's disease (AD) risk per *APOE* genotype was compared to the *APOE* $\epsilon 3/\epsilon 3$ reference group (i.e., odds ratio (OR) for *APOE* $\epsilon 3/\epsilon 3$ equals to 1), meta-analyzing results from the ADSP and ADGC cohorts (Stages 1+2). **eFigure 2** presents equivalent results at different inclusion cutoffs for European ancestry.