Supplementary material

Polygenic burden in focal and generalized epilepsies

Costin Leu, Remi Stevelink, Alexander Smith, Slavina B. Goleva, Masahiro Kanai, Lisa Ferguson, Ciaran Campbell, Yoichiro Kamatani, Yukinori Okada, Sanjay M. Sisodiya, Gianpiero L. Cavalleri, Bobby P.C. Koeleman, Holger Lerche, Lara Jehi, Lea K. Davis, Imad M. Najm, Aarno Palotie, Mark J. Daly, Robyn M. Busch, Epi25 Consortium, Dennis Lal

Corresponding author:

Dennis Lal, PhD

Genomic Medicine Institute

Lerner Research Institute

Cleveland Clinic

Cleveland, OH 44195, US

Email: lald@ccf.org

1. **Epi25 Consortium**
2. **Acknowledgments**
3. **Additional funding**
4. **Supplementary Methods**

4.1. Study cohorts
4.2. Cases
4.3. Controls
4.4. Biobank repositories
4.5. Data quality control and imputation
4.6. Detection of overlapping individuals across cohorts
4.7. Polygenic risk scoring SNP quality control
4.8. *P*-value thresholding

5. **Supplementary Figures**

5.1. Supplementary Figure 1: Principal component analysis of the European-ancestry study cohorts and 1000 Genomes Project samples
5.2. Supplementary Figure 2: Genome-wide polygenic risk for generalized epilepsy or focal epilepsy in the Finnish-ancestry population isolate
5.3. Supplementary Figure 3: Odds ratios for GE by GE-PRS deciles
5.4. Supplementary Figure 4: Odds ratios for FE by FE-PRS deciles

6. **Supplementary Tables**

6.1. Supplementary Table 1: GE- and FE-PRS in the European-ancestry exploration and replication cohorts
6.2. Supplementary Table 2: GE- and FE-PRS in the Finnish-ancestry population isolate
6.3. Supplementary Table 3: GE- and FE-PRS in the UKB, BioVU, and BBJ biobanks
6.4. Supplementary Table 4: *P*-value thresholding for GE-PRS
6.5. Supplementary Table 5: *P*-value thresholding for FE-PRS
6.6. Supplementary Table 6: Sample quality control filtering
6.7. Supplementary Table 7: Numbers of SNP considered for PRS
6.8. Supplementary Table 8: Diagnostic accuracy of the PRS in the *Epi25-EUR* and *Cleveland-EUR* cohorts

7. **References**

1. Epi25 Consortium

A full list of authors and affiliations

Epi25 sequencing, analysis, and project management at the Broad Institute

Yen-Chen Anne Feng[1,2,3,4], Daniel P. Howrigan[1,3,4], Liam E. Abbott[1,3,4], Katherine Tashman[1,3,4], Felecia Cerrato[3], Dennis Lal[4,5], Claire Churchhouse[1,3,4], Namrata Gupta[3], Benjamin M. Neale[1,3,4]

Epi25 executive committee

Samuel F. Berkovic[6], Holger Lerche[7], David B. Goldstein[8], Daniel H. Lowenstein[9]

Epi25 strategy, phenotyping, analysis, informatics, and project management committees

Samuel F. Berkovic[6], Holger Lerche[7], David B. Goldstein[8], Daniel H. Lowenstein[9], Gianpiero L. Cavalleri[10,11], Patrick Cossette[12], Chris Cotsapas[13], Peter De Jonghe[14,15,16], Tracy Dixon-Salazar[17], Renzo Guerrini[18], Hakon Hakonarson[19], Erin L. Heinzen[8], Ingo Helbig[20,21,22], Patrick Kwan[23,24], Anthony G. Marson[25], Slavé Petrovski[24,26], Sitharthan Kamalakaran[8], Sanjay M. Sisodiya[27], Randy Stewart[28], Sarah Weckhuysen[14,15,16], Chantal Depondt[29], Dennis J. Dlugos[19], Ingrid E. Scheffer[6,30], Pasquale Striano[31], Catharine Freyer[9], Roland Krause[32], Patrick May[32], Kevin McKenna[9], Brigid M. Regan[6], Susannah T. Bellows[6], Costin Leu[4,5,27]

Authors from individual Epi25 cohorts:

Australia: Melbourne (AUSAUS)

Samuel F. Berkovic[6], Ingrid E. Scheffer[6,30], Brigid M. Regan[6], Caitlin A. Bennett[6], Susannah T. Bellows[6], Esther M.C. Johns[6], Alexandra Macdonald[6], Hannah Shilling[6], Rosemary Burgess[6], Dorien Weckhuysen[6], Melanie Bahlo[33,34]

Australia: Royal Melbourne (AUSRMB)

Terence J. O'Brien[23,24], Patrick Kwan[23,24], Slavé Petrovski[24,26], Marian Todaro[23,24]

Belgium: Antwerp (BELATW)

Sarah Weckhuysen[14,15,16], Hannah Stamberger[14,15,16], Peter De Jonghe[14,15,16]

Belgium: Brussels (BELULB)

Chantal Depondt[29]

Canada: Andrade (CANUTN)

Danielle M. Andrade[35,36], Tara R. Sadoway[36], Kelly Mo[36]

Switzerland: Bern (CHEUBB)

Heinz Krestel[37], Sabina Gallati[38]

Cyprus (CYPCYP)

Savvas S. Papacostas[39], Ioanna Kousiappa[39], George A. Tanteles[40]

Czech Republic: Prague (CZEMTH)

Katalin Šterbová[41], Markéta Vlcková[42], Lucie Sedlácková[41], Petra Laššuthová[41]

Germany: Frankfurt/Marburg (DEUPUM)

Karl Martin Klein[43,44], Felix Rosenow[43,44], Philipp S. Reif[43,44], Susanne Knake[44]

Germany: Bonn (DEUUKB)

Wolfram S. Kunz[45,46], Gábor Zsurka[45,46], Christian E. Elger[46], Jürgen Bauer[46], Michael Rademacher[46]

Germany: Kiel (DEUUKL)

Ingo Helbig[20,21,22], Karl Martin Klein[43,44], Manuela Pendziwiat[21], Hiltrud Muhle[21], Annika Rademacher[21], Andreas van Baalen[21], Sarah von Spiczak[21], Ulrich Stephani[21], Zaid Afawi[47], Amos D. Korczyn[48], Moien Kanaan[49], Christina Canavati[49], Gerhard Kurlemann[50], Karen Müller-Schlüter[51], Gerhard Kluger[52,53], Martin Häusler[54], Ilan Blatt[48,55]

Germany: Leipzig (DEUULG)

Johannes R. Lemke[56], Ilona Krey[56]

Germany: Tuebingen (DEUUTB)

Holger Lerche[7], Yvonne G. Weber[7,57], Stefan Wolking[7], Felicitas Becker[7,58], Christian Hengsbach[7], Sarah Rau[7], Ana F. Maisch[7], Bernhard J. Steinhoff[59], Andreas Schulze-Bonhage[60], Susanne Schubert-Bast[61], Herbert Schreiber[62], Ingo Borggräfe[63], Christoph J. Schankin[64], Thomas Mayer[65], Rudolf Korinthenberg[66], Knut Brockmann[67], Gerhard Kurlemann[50], Dieter Dennig[68], Rene Madeleyn[69]

Finland: Kuopio (FINKPH)

Reetta Kälviäinen[70], Pia Auvinen[70], Anni Saarela[70]

Finland: Helsinki (FINUVH)

Tarja Linnankivi[71], Anna-Elina Lehesjoki[72]

Wales: Swansea (GBRSWU)

Mark I. Rees[73,74], Seo-Kyung Chung[73,74], William O. Pickrell[73], Robert Powell[73,75]

UK: UCL (GBRUCL)

Sanjay M. Sisodiya[27], Natascha Schneider[27], Simona Balestrini[27], Sara Zagaglia[27], Vera Braatz[27]

UK: Imperial/Liverpool (GBRUNL)

Anthony G. Marson[25], Michael R. Johnson[76], Pauls Auce[77], Graeme J. Sills[78]

Hong Kong (HKGHKK)

Patrick Kwan[23,24,79], Larry W. Baum[80,81,82], Pak C. Sham[80,81,82], Stacey S. Cherny[83], Colin H.T. Lui[84]

Croatia (HRVUZG)

Nina Barišic[85]

Ireland: Dublin (IRLRCI)

Gianpiero L. Cavalleri[10,11], Norman Delanty[10,11], Colin P. Doherty[86,11], Arif Shukralla[87], Mark McCormack[10], Hany El-Naggar[87,11]

Italy: Milan (ITAICB)

Laura Canafoglia[88], Silvana Franceschetti[88], Barbara Castellotti[89], Tiziana Granata[90]

Italy: Genova (ITAIGI)

Pasquale Striano[31], Federico Zara[91], Michele Iacomino[91], Francesca Madia[91], Maria Stella Vari[31], Maria Margherita Mancardi[91], Vincenzo Salpietro[31]

Italy: Bologna (ITAUBG)

Francesca Bisulli[92,93], Paolo Tinuper[92,93], Laura Licchetta[92,93], Tommaso Pippucci[94], Carlotta Stipa[92,93], Lorenzo Muccioli[92,93], Raffaella Minardi[92]

Italy: Catanzaro (ITAUMC)

Antonio Gambardella[95], Angelo Labate[95], Grazia Annesi[96], Lorella Manna[96], Monica Gagliardi[96]

Italy: Florence (ITAUMR)

Renzo Guerrini[18], Elena Parrini[18], Davide Mei[18], Annalisa Vetro[18], Claudia Bianchini[18], Martino Montomoli[18], Viola Doccini[18], Carla Marini[18]

Japan: RIKEN Institute (JPNRKI)

Toshimitsu Suzuki[97], Yushi Inoue[98], Kazuhiro Yamakawa[97]

Lithuania (LTUUHK)

Tumiene Birute[99], Mameniskiene Ruta[100], Utkus Algirdas[99], Praninskiene Ruta[100], Grikiniene Jurgita[100], Samaitiene Ruta[100]

New Zealand: Otago (NZLUTO)

Lynette G. Sadleir[101], Chontelle King[101], Emily Mountier[101]

Turkey: Bogazici (TURBZU)

S. Hande Caglayan[102], Mutluay Arslan[103], Zuhal Yapici[104], Uluc Yis[105], Pinar Topaloglu[104], Bulent Kara[106], Dilsad Turkdogan[107], Asli Gundogdu-Eken[102]

Turkey: Istanbul (TURIBU)

Nerses Bebek[108,109], Sibel Ugur-Iseri[109], Betül Baykan[108], Baris Salman[109], Garen Haryanyan[108], Emrah Yücesan[110], Yesim Kesim[108], Çigdem Özkara[111]

USA: BCH (USABCH)

Beth R. Sheidley[112,113], Catherine Shain[112,113], Annapurna Poduri[112,113]

USA: Philadelphia/CHOP (USACHP) and Philadelphia/Rowan (USACRW)

Russell J. Buono[114,115,19], Thomas N. Ferraro[114,22], Michael R. Sperling[115], Dennis J. Dlugos[19,22], Warren Lo[116], Michael Privitera[117], Jacqueline A. French[118], Patrick Cossette[12], Steven Schachter[119], Hakon Hakonarson[19]

USA: EPGP (USAEGP)

Daniel H. Lowenstein[9], Ruben I. Kuzniecky[120], Dennis J. Dlugos[19,22], Orrin Devinsky[118]

USA: NYU HEP (USAHEP)

Daniel H. Lowenstein[9], Ruben I. Kuzniecky[120], Jacqueline A. French[118], Manu Hegde[9]

USA: Penn/CHOP (USAUPN)

Ingo Helbig[20,22], Pouya Khankhanian[121,122], Katherine L. Helbig[20], Colin A. Ellis[122]

Epi25 control cohort

Gianfranco Spalletta[123,124], Fabrizio Piras[123], Federica Piras[123], Tommaso Gili[125,123], Valentina Ciullo[123,126]

Affiliations

1 Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
2 Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
3 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142, USA
4 Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
5 Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH 44195, USA
6 Epilepsy Research Centre, Department of Medicine, University of Melbourne, Victoria, Australia
7 Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany
8 Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA
9 Department of Neurology, University of California, San Francisco, CA 94110, USA

[10] The Department of Molecular and Cellular Therapeutics, The Royal College of Surgeons in Ireland, Dublin, Ireland
[11] The FutureNeuro Research Centre, Ireland
[12] University of Montreal, Montreal, QC H3T 1J4, Canada
[13] School of Medicine, Yale University, New Haven, CT 06510, USA
[14] Neurogenetics Group, University of Antwerp, Belgium
[15] Laboratory of Neurogenetics, Institute Born-Bunge, University of Antwerp, Belgium
[16] Division of Neurology, Antwerp University Hospital, Antwerp, Belgium
[17] LGS Foundation, NY 11716, USA
[18] Pediatric Neurology, Neurogenetics and Neurobiology Unit and Laboratories, Children's Hospital A. Meyer, University of Florence, Italy
[19] The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[20] Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[21] Department of Neuropediatrics, Christian-Albrechts-University of Kiel, 24105 Kiel, Germany
[22] Perelman School of Medicine, University of Pennsylvania, PA 19104, USA
[23] Department of Neuroscience, Central Clinical School, Monash University, Alfred Hospital, Melbourne, Australia
[24] Departments of Medicine and Neurology, University of Melbourne, Royal Melbourne Hospital, Parkville, Australia
[25] Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, UK
[26] Centre for Genomics Research, Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge, UK
[27] Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, UK and Chalfont Centre for Epilepsy, Chalfont St Peter, UK
[28] National Institute of Neurological Disorders and Stroke, MD 20852, USA
[29] Department of Neurology, Université Libre de Bruxelles, Brussels, Belgium
[30] Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Florey and Murdoch Children's Research Institutes, Melbourne, Australia
[31] Pediatric Neurology and Muscular Diseases Unit, Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, IRCCS "G. Gaslini" Institute, Genova, Italy
[32] Luxembourg Centre for Systems Biomedicine, University Luxembourg, Esch-sur-Alzette, Luxembourg
[33] Population Health and Immunity Division, the Walter and Eliza Hall Institute of Medical Research, Parkville 3052, VIC, Australia
[34] Department of Medical Biology, The University of Melbourne, Melbourne 3010, VIC, Australia
[35] Department of Neurology, Toronto Western Hospital, Toronto, ON M5T 2S8, Canada
[36] University Health Network, University of Toronto, Toronto, ON, Canada
[37] Departments of Neurology and BioMedical Research, Bern University Hospital and University of Bern, Bern, Switzerland
[38] Institute of Human Genetics, Bern University Hospital, Bern, Switzerland
[39] Neurology Clinic B, The Cyprus Institute of Neurology and Genetics, 2370 Nicosia, Cyprus

[40] Department of Clinical Genetics, The Cyprus Institute of Neurology and Genetics, 2370 Nicosia, Cyprus
[41] Department of Paediatric Neurology, 2nd Faculty of Medicine, Charles University and Motol Hospital, Prague, Czech Republic
[42] Department of Biology and Medical Genetics, 2nd Faculty of Medicine, Charles University and Motol Hospital, Prague, Czech Republic
[43] Epilepsy Center Frankfurt Rhine-Main, Center of Neurology and Neurosurgery, Goethe University Frankfurt, Frankfurt, Germany
[44] Epilepsy Center Hessen-Marburg, Department of Neurology, Philipps University Marburg, Marburg, Germany
[45] Institute of Experimental Epileptology and Cognition Research, University Bonn, 53127 Bonn, Germany
[46] Department of Epileptology, University Bonn, 53127 Bonn, Germany
[47] Sackler School of Medicine, Tel-Aviv University, Ramat Aviv, Israel
[48] Tel-Aviv University Sackler Faculty of Medicine, Ramat Aviv 69978, Israel
[49] Hereditary Research Lab, Bethlehem University, Bethlehem, Palestine
[50] Department of Neuropediatrics, Westfälische Wilhelms-University, Münster, Germany
[51] Epilepsy Center for Children, University Hospital Neuruppin, Brandenburg Medical School, Neuruppin, Germany
[52] Neuropediatric Clinic and Clinic for Neurorehabilitation, Epilepsy Center for Children and Adolescents, Vogtareuth, Germany
[53] Research Institute Rehabilitation / Transition / Palliation, PMU Salzburg, Austria
[54] Division of Neuropediatrics and Social Pediatrics, Department of Pediatrics, University Hospital, RWTH Aachen, Aachen, Germany
[55] Department of Neurology, Sheba Medical Center, Ramat Gan, Israel
[56] Institute of Human Genetics, Leipzig, Germany
[57] Department of Neurosurgery, University of Tübingen, Germany
[58] RKU-University Neurology Clinic of Ulm, Ulm, Germany
[59] Kork Epilepsy Center, Kehl-Kork, Germany
[60] Epilepsy Center, University of Freiburg, Freiburg im Breisgau, Germany
[61] Section Neuropediatrics and Inborn Errors of Metabolism, University Children's Hospital, Heidelberg, Germany
[62] Neurological Practice Center & Neuropoint Patient Academy, Ulm, Germany
[63] Department of Pediatric Neurology and Developmental Medicine, LMU Munich, Munich, Germany
[64] Department of Neurology, University of Munich Hospital-Großhadern, Munich, Germany
[65] Saxonian Epilepsy Center Radeberg, Radeberg, Germany
[66] Division of Neuropediatrics and Muscular Disorders, University Hospital Freiburg, Freiburg, Germany
[67] University Children's Hospital, Göttingen, Germany
[68] Private Neurological Practice, Stuttgart, Germany
[69] Department of Pediatrics, Filderklinik, Filderstadt, Germany
[70] Neurocenter, Kuopio University Hospital, Kuopio Finland and Institute of Clinical Medicine, University of Eastern Finland, Finland
[71] Child Neurology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[72] Medicum, University of Helsinki, Helsinki, Finland, and Folkhälsan Research Center, Helsinki, Finland
[73] Neurology Research Group, Swansea University Medical School, Swansea University SA2 8PP, UK
[74] Faculty of Medicine and Health, University of Sydney, Sydney, Australia
[75] Department of Neurology, Morriston Hospital, Abertawe Bro Morgannwg Health University Board, Swansea, UK
[76] Division of Brain Sciences, Imperial College London, London, UK
[77] Department of Neurology, Walton Centre NHS Foundation Trust, Liverpool, UK
[78] School of Life Sciences, University of Glasgow, Glasgow, UK
[79] Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong, China
[80] The State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong, China
[81] Centre for Genomic Sciences, University of Hong Kong, Hong Kong, China
[82] Department of Psychiatry, University of Hong Kong, Hong Kong, China
[83] Department of Epidemiology and Preventive Medicine and Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Israel
[84] Department of Medicine, Tseung Kwan O Hospital, Hong Kong, China
[85] Department of Pediatric University Hospital center Zagreb, Croatia
[86] Neurology Department, St. James Hospital, Dublin, Ireland
[87] The Department of Neurology, Beaumont Hospital, Dublin, Ireland
[88] Neurophysiopathology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[89] Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[90] Department of Pediatric Neuroscience, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[91] Laboratory of Neurogenetics, IRCCS "G. Gaslini" Institute, Genova, Italy
[92] IRCCS, Institute of Neurological Sciences of Bologna, Bologna, Italy
[93] Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy
[94] Medical Genetics Unit, Polyclinic Sant'Orsola-Malpighi University Hospital, Bologna, Italy
[95] Institute of Neurology, Department of Medical and Surgical Sciences, University "Magna Graecia", Catanzaro, Italy
[96] Institute of Molecular Bioimaging and Physiology, CNR, Section of Germaneto, Catanzaro, Italy
[97] Laboratory for Neurogenetics, RIKEN Center for Brain Science, Saitama, Japan
[98] National Epilepsy Center, Shizuoka Institute of Epilepsy and Neurological Disorder, Shizuoka, Japan
[99] Institute of Biomedical Sciences, Faculty of Medicine, Vilnius University, Vilnius, Lithuania
[100] Institute of Clinical Medicine, Faculty of Medicine, Vilnius University, Vilnius, Lithuania
[101] Department of Paediatrics and Child Health, University of Otago, Wellington, New Zealand
[102] Department of Molecular Biology and Genetics, Bogaziçi University, Istanbul, Turkey
[103] Department of Child Neurology, Gulhane Education and Research Hospital, Health Sciences University, Ankara, Turkey
[104] Department of Child Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey

[105] Department of Child Neurology, Medical School, Dokuz Eylul University, Izmir, Turkey
[106] Department of Child Neurology, Medical School, Kocaeli University, Kocaeli, Turkey
[107] Department of Child Neurology, Medical School, Marmara University, Istanbul, Turkey
[108] Department of Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey
[109] Department of Genetics, Aziz Sancar Institute of Experimental Medicine, Istanbul University, Istanbul, Turkey
[110] Bezmialem Vakif University, Institute of Life Sciences and Biotechnology, Istanbul, Turkey
[111] Department of Neurology, Faculty of Medicine, Cerrahpasa University Istanbul, Istanbul, Turkey
[112] Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, MA 02115, USA
[113] Department of Neurology, Harvard Medical School, Boston, MA 02115, USA
[114] Cooper Medical School of Rowan University, Camden, NJ 08103, USA
[115] Thomas Jefferson University, Philadelphia, PA 19107, USA
[116] Nationwide Children's Hospital, Columbus, OH 43205, USA
[117] University of Cincinnati, Cincinnati, OH 45220, USA
[118] Department of Neurology, New York University/Langone Health, New York, NY 10016, USA
[119] Beth Israel Deaconess/Harvard, Boston, MA 02115, USA
[120] Department of Neurology, Hofstra-Northwell Medical School, New York, NY 11549, USA
[121] Center for Neuro-engineering and Therapeutics, University of Pennsylvania, Philadelphia, PA 19104, USA
[122] Department of Neurology, Hospital of University of Pennsylvania, Philadelphia, PA 19104, USA
[123] Neuropsychiatry Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy
[124] Division of Neuropsychiatry, Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA
[125] IMT School for Advanced Studies Lucca, Lucca, Italy
[126] Department of Neurosciences, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy

2. Acknowledgments

We gratefully thank the Epi25 principal investigators, local staff from individual cohorts, and all the epilepsy patients who participated in the study for making possible this global collaboration and resource to advance epilepsy genetics research. This work is part of the Centers for Common Disease Genomics (CCDG) program, funded by the National Human Genome Research Institute (NHGRI) and the National Heart, Lung, and Blood Institute (NHLBI). CCDG-funded Epi25 research activities at the Broad Institute, including genomic data generation in the Broad Genomics Platform, are supported by NHGRI grant UM1 HG008895 (PIs: Eric Lander, Stacey

Gabriel, Mark Daly, Sekar Kathiresan). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We are grateful for support from the Epilepsy Society, UK, and the Muir Maxwell Trust. Part of this work was undertaken at University College London Hospitals, which received a proportion of funding from the NIHR Biomedical Research Centres Funding Scheme. The Vanderbilt University Medical Center's biobank (BioVU) was supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. This research has been conducted using the UK Biobank Resource under Application Number '35124'. This work involved the use of the Enterprise Research Infrastructure & Services (ERIS) at Partners HealthCare. We thank the Partners HealthCare Biobank for providing genomic and health information data. The in-house project on inflammatory bowel disease without reported epilepsy was generated as part of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) IBD Genetics Consortium (IBDGC) supported by the Helmsley Charitable Trust and the Centers for Common Disease Genomes Program (NHGRI). DNA samples were obtained from the following collections: The Lunenfeld-Tanenbaum Research Institute Mount Sinai Hospital (PI: Mark Silverberg), The University of Pittsburgh School of Medicine (PI: Richard Duerr), The Emory University School of Medicine (PI: Subra Kugathasan), The Johns Hopkins Hospital (PI: Steven Brant), The Icahn School of Medicine at Mount Sinai (PI: Judy Cho), The Washington University School of Medicine (PI: Rodney Newberry), The University of Miami Miller School of Medicine (PI: Maria Abreu, Jake McCauley), and Cedars Sinai (PI: Dermot McGovern, Stephan Targan).

For the FINRISK population controls, we thank: Aarno Palotie[1,2,3,4,5]

Affiliations

1 Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
2 Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
3 Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
4 Institute for Molecular Medicine Finland, University of Helsinki, Helsinki 00014, Finland
5 Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA

For the Genomic Psychiatry Cohort (GPC) healthy controls, we thank: Michele T. Pato[1], Carlos N. Pato[1], Evelyn J. Bromet[2], Celia Barreto Carvalho[3], Eric D. Achtyes[4], Maria Helena Azevedo[5], Roman Kotov[2], Douglas S. Lehrer[6], Dolores Malaspina[7], Stephen R. Marder[8], Helena Medeiros[1], Christopher P. Morley[9], Diana O. Perkins[10], Janet L Sobell[11], Peter F. Buckley[12], Fabio Macciardi[13], Mark H. Rapaport[14], James A. Knowles[1], Genomic Psychiatry Cohort (GPC) Consortium, Ayman H. Fanous[1,15], Steven A. McCarrol[16,17,18]

Affiliations

1 Department of Psychiatry and Behavioral Sciences, SUNY Downstate Medical Center, Brooklyn, NY, USA
2 Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA
3 Faculty of Social and Human Sciences, University of Azores, PT
4 Cherry Health and Michigan State University College of Human Medicine, Grand Rapids, MI, USA
5 Institute of Medical Psychology, Faculty of Medicine, University of Coimbra, Coimbra, PT
6 Department of Psychiatry, Wright State University, Dayton, OH, USA
7 Departments of Psychiatry, Genetics & Genomics, Icahn School of Medicine at Mount Sinai, NY, USA
8 Semel Institute for Neuroscience at UCLA, Los Angeles, CA, USA
9 Departments of Public Health and Preventive Medicine, Family Medicine, and Psychiatry and Behavioral Sciences, State University of New York, Upstate Medical University, Syracuse, NY, USA
10 Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA
11 Department of Psychiatry & Behavioral Sciences, University of Southern California, Los Angeles, CA, USA
12 Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, VA, USA
13 Department of Psychiatry and Human Behavior, University of California, Irvine, CA, USA
14 Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA
15 Department of Psychiatry, Veterans Administration New York Harbor Healthcare System, Brooklyn, NY, USA
16 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142, USA
17 Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
18 Department of Genetics, Harvard Medical School, Boston, MA, 02115, USA

3. Additional funding

SMS received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 279062 (EpiPGX) and the Department of Health's NIHR

Biomedical Research Centers funding scheme, and the UK Epilepsy Society. RMB received support from the NIH/NCATS, CTSA UL1TR000439, Cleveland, Ohio.

4. Supplementary Methods

4.1. Study cohorts

Patients or their legal guardians provided signed informed consent according to local IRB requirements. Samples had been collected over 20 years in some centers, so the consent forms reflected standards at the time of collection. For Epi25 Consortium samples collected after January 25, 2015, consent forms required specific language according to the NIH Genomic Data Sharing Policy (https://osp.od.nih.gov/wp-content/uploads/NIH_GDS_Policy.pdf).

4.2. Cases

We used three independent patient cohorts, each consisting of individuals with generalized epilepsy (GE) or focal epilepsy (FE). The first cohort was derived from the European-ancestry subsample of the Epi25 project, an international multi-center epilepsy genetics research consortium, comprising 5,705 people with GE (*GE-Epi25-EUR* cohort) or FE (*FE-Epi25-EUR* cohort) after quality control (QC). The second cohort was derived from a single clinical center, the Cleveland Clinic Epilepsy Center, comprising 620 people with GE (*GE-Cleveland-EUR*) or FE (*FE-Cleveland-EUR*) after QC, all of European ancestry. The third cohort was derived from the Finnish-ancestry subsample of the Epi25 project that was not part of the first cohort, comprising 449 people with GE (*GE-Epi25-FIN*) or FE (*FE-Epi25-FIN*) after QC. All cohorts are detailed in Table 1. All three patient cohorts were genotyped with Illumina's Global Screening Array. GE and FE were diagnosed in all cohorts according to clinical criteria (clinical interview, neurological examination, EEG, imaging data). International League Against Epilepsy (ILAE) classifications (Berg *et al.*, 2010) were strictly followed in the Epi25 cohorts.

4.3. Controls

The European-ancestry epilepsy cohorts were matched to population controls from four merged cohorts: (1) healthy individuals from the Epi25 project (N = 210); (2) an in-house project on inflammatory bowel disease without reported epilepsy (N = 4,905); (3) healthy individuals from the Genetics of Personality Consortium (N = 463); and (4) population controls without reported epilepsy or potentially epileptogenic brain diseases (G40: epilepsy and recurrent seizures, C71: malignant neoplasm of brain, F44.5: conversion disorder with seizures or convulsions, G81.0:

flaccid hemiplegia, G93: other disorders of brain, I61: nontraumatic intracerebral hemorrhage,, I67: other cerebrovascular diseases, P90: convulsions of newborn, R56.0: febrile convulsions) from the Partners HealthCare Biobank (N = 14,875) (Karlson *et al.*, 2016). The first three control cohorts were genotyped with Illumina's Global Screening Array. The fourth control cohort was genotyped with Illumina's Multi-Ethnic Global Screening Array. The merged population cohort comprised 20,435 controls after QC. Population controls for the Finnish-ancestry cohort were obtained from the THL Institute for Health and Welfare (subsample of the FINRISK study, N = 1,559) (Borodulin *et al.*, 2017). The Finnish-ancestry cohort was genotyped on the Illumina's Global Screening Array.

4.4. Biobank repositories

Three large-scale biobank repositories were available as additional replication cohorts: UK Biobank (UKB), N = 383,656 (Sudlow *et al.*, 2015); Vanderbilt University biorepository (BioVU), N = 49,494 (Roden *et al.*, 2008); BioBank Japan (BBJ), N = 168,680 (Nagai *et al.*, 2017). Seizure or epilepsy classification was available as International Classification of Diseases (ICD-10) codes. We used ICD-10 G40.3 codes to identify people with GE, and G40.0 to G40.2 codes to identify people with FE. To increase the phenotypic homogeneity of each group, we excluded people with ICD-10 codes for both seizure types. In addition, we excluded subjects with epilepsy and other potentially epileptogenic brain diseases (ICD-10 C71: malignant neoplasm of brain, I61: nontraumatic intracerebral hemorrhage, I67: other cerebrovascular diseases, G81: hemiplegia and hemiparesis, G93: other disorders of brain, Q28: other congenital malformations of circulatory system).

4.5. Data quality control and imputation

Before imputation, we excluded genotyped individuals based on the following criteria: (1) genotype call rate < 0.95; (2) high (> 0.2) or low (< -0.2) inbreeding coefficient estimate of the observed versus expected number of homozygous genotypes; (3) missing, ambiguous, or sex mismatch between X-chromosome genotype and reported gender; (4) population outliers not clustering with the 1000 Genomes Project (1000 Genomes Project Consortium *et al.*, 2015) European samples in a principal component analysis (Supplementary Fig. 1). Then, we excluded single-nucleotide polymorphisms (SNPs) based on the following criteria: (1) SNP call rate < 0.98 in the combined case/control dataset; (2) minor allele frequency (MAF) < 0.01; (3) deviation from the Hardy-Weinberg equilibrium with $P < 1 \times 10^{-6}$. Sample and SNP QC

procedures were performed using PLINK v1.9 (Chang *et al.*, 2015) and GCTA (Yang *et al.*, 2011). The genotyped dataset was aligned to the imputation reference (variant name, variant position, and strand orientation) using the Genotype Harmonizer (Deelen *et al.*, 2014). Imputation to the Haplotype Reference Consortium (HRC) reference r1.1 (McCarthy *et al.*, 2016) was performed using reference-based phasing with Eagle v2.4 (Loh *et al.*, 2016) and Minimac4 (https://github.com/statgen/Minimac4), as implemented on the Michigan Imputation Server (Das *et al.*, 2016). After imputation, we removed randomly one individual from each pair of individuals with 3rd-degree relationships and higher (kinship coefficient > 0.0442) using KING (Manichaikul *et al.*, 2010). Detailed information on all excluded individuals is given in Supplementary Table 6.

4.6. Detection of overlapping individuals across cohorts

A proportion of cases and controls who were part of the prior GWAS (International League Against Epilepsy Consortium on Complex Epilepsies, 2018) and thus contributed to PRS development, were also genotyped in our study cohorts. These individuals were excluded from the study cohorts. Inspired by the one-way cryptographic hash function (Turchin and Hirschhorn, 2012), we used a protocol that allows overlapping participants between studies to be identified without sharing individual-level data. One-way cryptographic hashes are a form of security algorithms that alter input data in such a way that the resulting output data cannot be reverted feasibly to the original form. To identify overlapping individuals, we first generated ten batches of SNPs, which did not have missing genotypes in any individual in this study or the GWAS used as the source for the generation of the PRS (N_{SNP} = 25). We then computed hash values (checksums) for each of the ten batches for each individual, using the "cksum" command, which is routinely available in Linux operating systems. The "cksum" command will always generate the same unique hash value for each batch, when using the same SNPs, with the same information (same non-missing genotype), and in the same order (sorted by physical position). We then marked every pair of individuals with one or more identical hash values (out of the ten) as duplicate and excluded the corresponding individual from our study. The procedure is implemented in perl and freely available at https://personal.broadinstitute.org/sripke/share_links/checksums_download/.

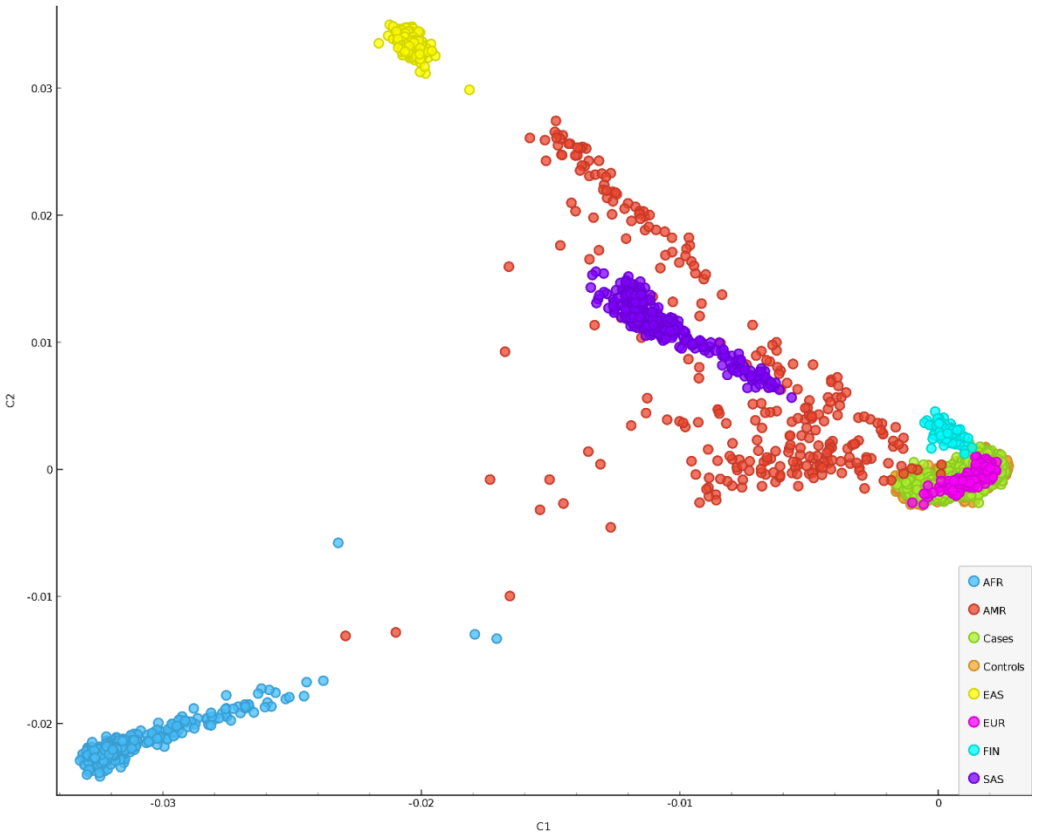
4.7. Polygenic risk scoring SNP quality control

We selected high-quality imputed and genotyped SNPs in the study cohorts based on the following criteria and established best practices (Choi *et al.*, 2018): (1) Minimac4 imputation quality score, $R^2 > 0.3$; (2) Minimac4 squared correlation value between masked genotypes of genotyped SNPs and the imputed dosages, Emp-$R^2 > 0.3$; (3) call rate > 0.98 in either cases or controls; (4) minor allele frequency > 2% in either cases or controls; (5) deviation from the Hardy-Weinberg equilibrium with $P > 10^{-7}$ in either cases or controls; (6) SNPs with non-ambiguous alleles (A/T or C/G excluded). We generated PRS based on the overlap of the remaining SNPs and GWAS SNPs with available summary statistics, pruned to a subset of uncorrelated SNPs ($r^2 < 0.1$ within 500kb from the most significant SNP in each locus). The numbers of SNPs available for PRS generation are detailed in Supplementary Table 7.

4.8. *P*-value thresholding

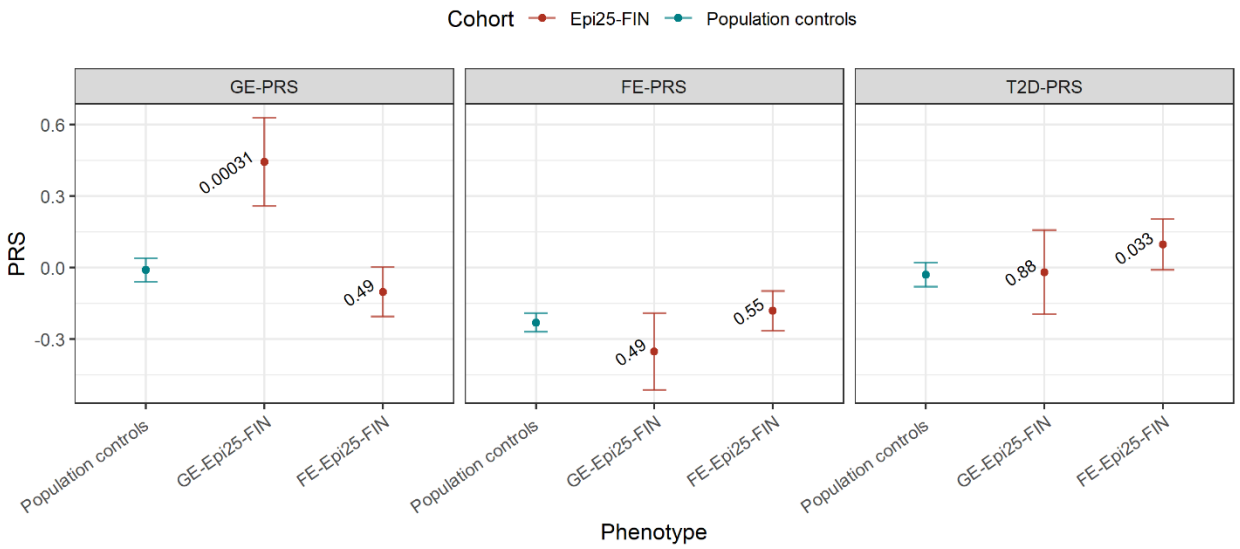
To identify the optimal *P*-value threshold for PRS prediction, we performed a random split of our exploration cohort (Epi25-EUR, 5,705 European-ancestry individuals: 80% training, 20% validation), and generated GE-PRS (Supplementary Table 4) and FE-PRS (Supplementary Table 5) at eight *P*-value thresholds (PT = 10^{-4}, 10^{-3}, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) in the corresponding training samples. The performance of the best predicting threshold was explored in the validation samples. GE-PRS predicted the GE vs. control status best at PT=0.5 in the training sample. The threshold was confirmed in the validation sample by a significant prediction and a similar level of the explained phenotypic variance (training: 2.79% / validation: 2.86%). FE-PRS predicted the FE vs. control status best at PT=0.1. In the validation sample we observed a slight loss of power at PT=0.1, when considering the explained phenotypic variance (training: 0.62% / validation: 0.28%). For a homogeneous methodology for the whole study, we also explored PT=0.5 in the FE validation sample and observed a more significant prediction and a full recovery of the explained phenotypic variance (training: 0.56% / validation: 0.52%). Subsequently, PT=0.5 was considered as the optimal PT for FE, because of negligible differences of the prediction power at all PTs from 0.1 to 0.5 in the training sample. The identified optimal *P*-value threshold (0.5 for GE and FE) was applied for all datasets.

5. Supplementary Figures



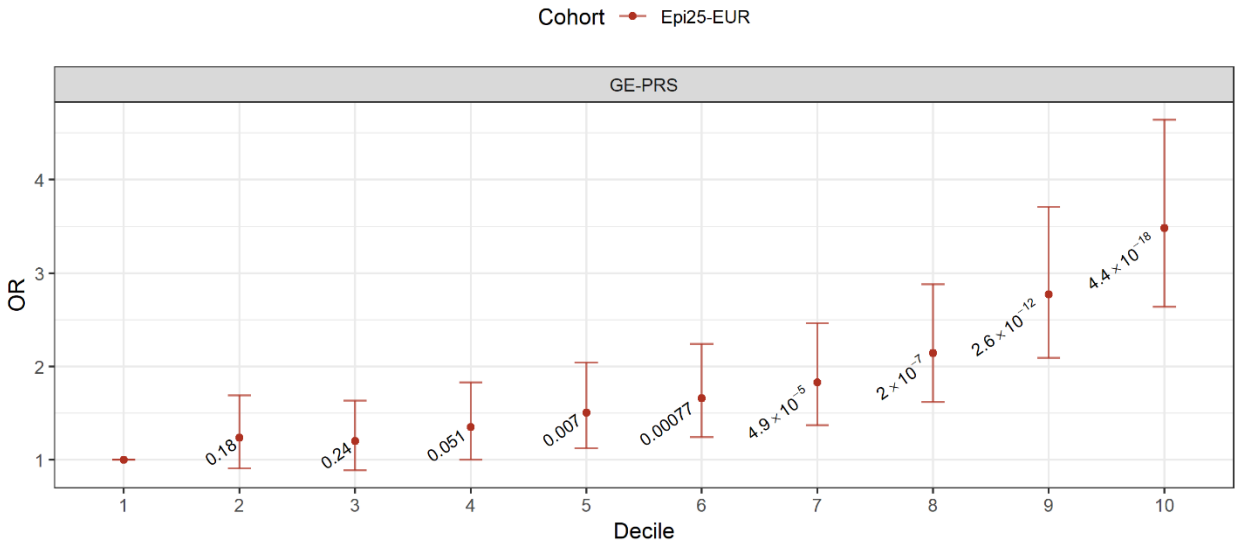
Supplementary Figure 1: Principal component analysis of the European-ancestry study cohorts and 1000 Genomes Project samples

Legend: C1: PCA principal component 1, C2: PCA principal component 2, AFR: 1000 Genomes African samples, AMR: 1000 Genomes admixed American samples, Cases: European Epi25 Consortium and Cleveland Clinic samples, Controls: European merged population samples, EAS: 1000 Genomes East Asian samples, EUR: 1000 Genomes European samples, FIN: 1000 Genomes South Finnish samples, SAS: 1000 Genomes South Asian samples.



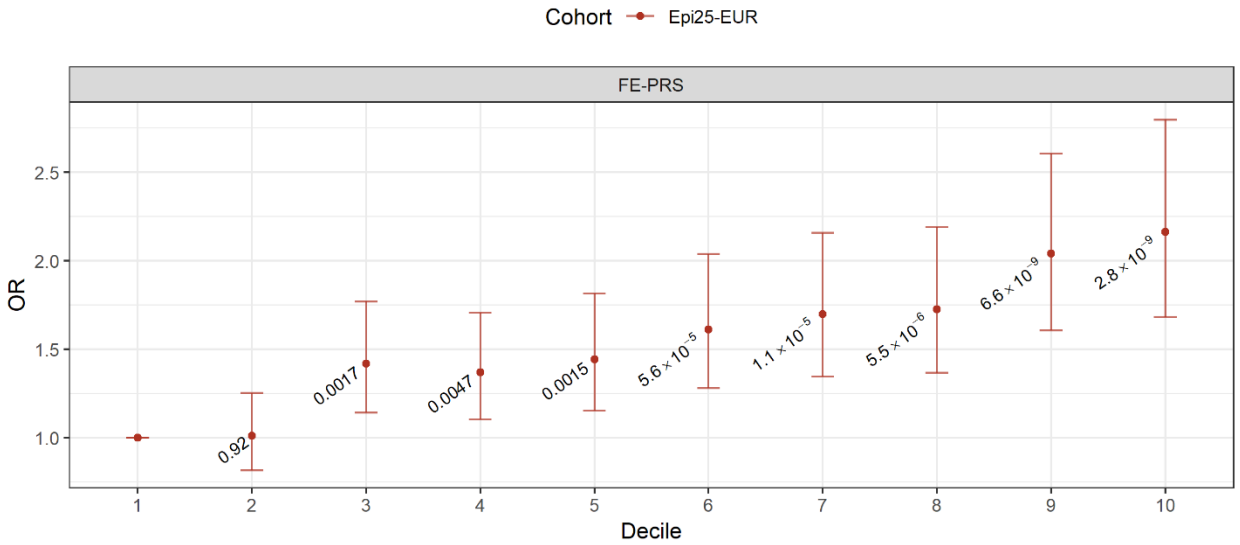
Supplementary Figure 2: Genome-wide polygenic risk for generalized epilepsy or focal epilepsy in the Finnish-ancestry population isolate

Shown are the means of the standardized GE-, FE-, and T2D-PRS with 95% confidence intervals for the Finnish-ancestry population controls (highlighted in blue; N = 1,559) and the Finnish-ancestry generalized epilepsy and focal epilepsy Epi25 cohorts (highlighted in red; *GE-Epi25-FIN* N = 112; *FE-Epi25-FIN* N = 337). The *P*-values for the differences between cases and population controls are given as numbers. The threshold for statistical significance after Bonferroni correction was set to $\alpha = 1.67 \times 10^{-2}$ (three tests per cohort).



Supplementary Figure 3: Odds ratios for GE by GE-PRS deciles

Based on the GE-PRS, individuals with GE and controls were allocated to 10 deciles containing near identical numbers of individuals. Decile 1 contained the lowest scores and was used as reference for deciles 2-10 that had increasingly higher GE-PRS. Odds ratios and *P*-values were calculated using a logistic regression model of deciles 2-10 against decile 1, adjusted for sex and the first four principal components of ancestry. The points represent the odds ratios. The bars represent the lower and upper confidence intervals of the odds ratios.



Supplementary Figure 4: Odds ratios for FE by FE-PRS deciles

Based on the FE-PRS, individuals with FE and controls were allocated to 10 deciles containing near identical numbers of individuals. Decile 1 contained the lowest scores and was used as reference for deciles 2-10 that had increasingly higher FE-PRS. Odds ratios and *P*-values were calculated using a logistic regression model of deciles 2-10 against decile 1, adjusted for sex and the first four principal components of ancestry. The points represent the odds ratios. The bars represent the lower and upper confidence intervals of the odds ratios.

6. Supplementary Tables

Supplementary Table 1: GE- and FE-PRS in the European-ancestry exploration and replication cohorts

					GE-PRS (N_{SNP} = 50,515)		FE-PRS (N_{SNP} = 51,333)		T2D-PRS (N_{SNP} = 72,305)	
Center/Study	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value**	**% R^2**	***P*-value**	**% R^2**	***P*-value**	**% R^2**
Epi25-EUR	EUR	GE	2,256	20,435	2.35E-70	2.83	1.71E-15	0.56	0.61	0.002
Cleveland-EUR	EUR	GE	85	20,435	1.43E-07	2.58	4.72E-03	0.74	0.92	0.001
Epi25-EUR	EUR	FE	3,449	20,435	8.21E-18	0.51	5.74E-19	0.55	8.48E-03	0.05
Cleveland-EUR	EUR	FE	535	20,435	6.12E-04	0.26	1.69E-06	0.51	0.017	0.13
Epi25-EUR	EUR	GE vs. FE	2,256	3,449	1.64E-15	1.74	0.84	0.001	0.40	0.02
Cleveland-EUR	EUR	GE vs. FE	85	535	2.85E-04	3.87	0.42	0.19	0.30	0.31

P-values were calculated using a logistic regression model, adjusted for sex and the first four principal components of ancestry. The threshold for statistical significance after Bonferroni correction was set to $\alpha = 1.67 \times 10^{-2}$ (three tests per cohort). Legend: PRS: polygenic risk score, N: number, SNP: single nucleotide polymorphism, GE: generalized epilepsy, FE: focal epilepsy, T2D: Type 2 diabetes, EUR: European, % R^2: percentage of explained variance (Nagelkerke's pseudo-R^2).

Supplementary Table 2: GE- and FE-PRS in the Finnish-ancestry population isolate

					GE-PRS (N_{SNP} = 59,006)		FE-PRS (N_{SNP} = 59,728)		T2D-PRS (N_{SNP} = 110,915)	
Center/Study	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value**	**% R^2**	***P*-value**	**% R^2**	***P*-value**	**% R^2**
Epi25-FIN	FIN	GE	112	1559	3.11E-04	2.01	0.49	0.04	0.88	0.003
Epi25-FIN	FIN	FE	337	1559	0.49	0.04	0.55	0.03	0.033	0.39
Epi25-FIN	FIN	GE vs. FE	112	337	1.80E-04	4.58	0.35	0.27	0.36	0.26

P-values were calculated using a logistic regression model, adjusted for sex and the first four principal components of ancestry. The threshold for statistical significance after Bonferroni correction was set to $\alpha = 1.67 \times 10^{-2}$ (three tests per cohort). Legend: PRS: polygenic risk score, N: number, SNP: single nucleotide polymorphism, GE: generalized epilepsy, FE: focal epilepsy, T2D: Type 2 diabetes, FIN: Finnish, % R^2: percentage of explained variance (Nagelkerke's pseudo-R^2).

Supplementary Table 3: GE- and FE-PRS in the UKB, BioVU, and BBJ biobanks

UK Biobank					**GE-PRS (N_{SNP} = 62,986)**		**FE-PRS (N_{SNP} = 63,730)**		**T2D-PRS (N_{SNP} = 143,464)**	
Center/Study	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value**	**% R^2**	***P*-value**	**% R^2**	***P*-value**	**% R^2**
UKB	EUR	GE	246	383,197	2.89E-02	0.12	0.069	0.08	0.88	0.001
UKB	EUR	FE	213	383,197	0.063	0.10	0.44	0.02	0.84	0.001
Vanderbilt biorepository					**GE-PRS (N_{SNP} = 87,758)**		**FE-PRS (N_{SNP} = 88,468)**		**T2D-PRS (N_{SNP} = 246,721)**	
Center/Study	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value**	**% R^2**	***P*-value**	**% R^2**	***P*-value**	**% R^2**
BioVU	EUR	GE	293	48,665	1.09E-02	0.19	0.88	0.001	0.85	0.001
BioVU	EUR	FE	536	48,665	0.37	0.01	0.23	0.03	0.61	0.005
BioBank Japan					**GE-PRS (N_{SNP} = 52,021)**		**FE-PRS (N_{SNP} = 52,504)**		**T2D-PRS (N_{SNP} = 65,379)**	
Center/Study	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value**	**% R^2**	***P*-value**	**% R^2**	***P*-value**	**% R^2**
BBJ	JPN	GE	219	168,356	0.33	0.03	0.32	0.03	0.45	0.02
BBJ	JPN	FE	105	168,356	0.55	0.02	0.29	0.06	0.096	0.16

P-values were calculated using a logistic regression model adjusted for sex and the first four principal components of ancestry. The threshold for statistical significance after Bonferroni correction was set to $\alpha = 1.67 \times 10^{-2}$ (two tests per cohort and one meta-analysis). Legend: PRS: polygenic risk score, UKB: UK Biobank, BioVU: Vanderbilt University biorepository, BBJ: BioBank Japan, N: number, SNP: single nucleotide polymorphism, GE: generalized epilepsy, FE: focal epilepsy, T2D: Type 2 diabetes, EUR: European, JPN: Japanese, % R^2: percentage of explained variance (Nagelkerke's pseudo-R^2).

Supplementary Table 4: *P*-value thresholding for GE-PRS

Training set	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value threshold**	***P*-value**	**% R^2**	**N_{SNP}**
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.0001	5.89E-20	0.93	165
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.001	1.32E-36	1.78	772
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.05	1.57E-52	2.61	11,562
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.1	4.83E-55	2.74	19,032
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.2	6.22E-55	2.74	29,620
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.3	4.10E-54	2.70	38,005
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.4	4.24E-55	2.75	44,794
Epi25-EUR (80% samples)	EUR	GE	1,805	16,348	0.5	8.30E-56	2.79	50,530
Validation set	Ethnicity	Epilepsy	Cases, N	Controls, N	*P*-value threshold	*P*-value	% R^2	N_{SNP}
Epi25-EUR (20% samples)	EUR	GE	451	4,087	0.5	1.32E-15	2.86	50,482

Eight *P*-value thresholds (10^{-4}, 10^{-3}, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) were tested in a random split of the GE exploration cohort (*GE-Epi25-EUR*) into training (80%) and validation (20%). The optimal *P*-value threshold is highlighted with a blue border. Legend: N: number, % R^2: percentage of explained variance (Nagelkerke's pseudo-R^2), SNP: single nucleotide polymorphism, GE: generalized epilepsy.

Supplementary Table 5: *P*-value thresholding for FE-PRS

Training set	Ethnicity	Epilepsy	Cases, N	Controls, N	*P*-value threshold	*P*-value	% R^2	N_{SNP}
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.0001	0.78	6.52E-04	117
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.001	1.01E-03	0.09	694
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.05	1.33E-13	0.48	11,721
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.1	3.74E-17	0.62	19,394
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.2	1.88E-15	0.55	30,071
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.3	2.94E-16	0.58	38,499
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.4	1.12E-15	0.56	45,571
Epi25-EUR (80% samples)	EUR	FE	2,760	16,348	0.5	7.80E-16	0.56	51,336
Validation set	**Ethnicity**	**Epilepsy**	**Cases, N**	**Controls, N**	***P*-value threshold**	***P*-value**	**% R^2**	**N_{SNP}**
Epi25-EUR (20% samples)	EUR	FE	689	4,087	0.1	4.36E-03	0.28	19,393
Epi25-EUR (20% samples)	EUR	FE	689	4,087	0.5	1.10E-04	0.52	51,321

Eight *P*-value thresholds (10^{-4}, 10^{-3}, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) were tested in a random split of the FE exploration cohort (*FE-Epi25-EUR*) into training (80%) and validation (20%). The optimal *P*-value thresholds are highlighted with blue borders. Legend: N: number, % R^2: percentage of explained variance (Nagelkerke's pseudo-R^2), SNP: single nucleotide polymorphism, FE: focal epilepsy.

Supplementary Table 6: Sample quality control filtering

Cohort	Epi25		Cleveland		
Phenotype	**GE**	**FE**	**GE**	**FE**	**Controls**
Available samples before quality control	4,613	6,809	97	633	28,187
Low genotyping call rate < 0.95	-18	-11	0	0	-8
High (> 0.2) or low (< -0.2) inbreeding coefficient estimate	-21	-46	0	0	-67
Missing sex	-10	-7	0	-1	-37
Ambiguous sex (undeterminable by genotypes)	-10	-43	0	0	-3
Mismatch between genotyped and reported gender	-102	-99	0	0	-154
Non-European individuals	-791	-752	-11	-92	-5026
Imputation	3,661	5,851	86	540	22,892
Related individuals across study cohorts	-114	-115	0	-1	-809
Overlapping individuals across study cohorts and GWAS	-1179	-1950	-1	-4	-89
Samples available for PRS after quality control	2,368	3,786	85	535	21,994

Overview of the Epi25 (EUR and FIN) and Cleveland Clinic cohorts before and after quality control filtering for PRS generation.

Supplementary Table 7: Numbers of SNP considered for PRS

	N	
Imputed SNPs before QC	30,260,497	
SNPs after post-imputation QC (Supplementary material 4.7)	2,194,578	
SNPs for PRS after P-value thresholding and LD-pruning (Supplementary material 4.7 and 4.8)	N	mean absolute Beta
GE-PRS in EUR samples	50,515	0.0073
FE-PRS in EUR samples	51,333	0.0090
T2D-PRS in EUR samples	72,305	0.0012
GE-PRS in FIN samples	59,006	0.0074
FE-PRS in FIN samples	59,728	0.0093
T2D-PRS in FIN samples	110,915	0.0017

Supplementary Table 8: Diagnostic accuracy of the PRS in the *Epi25-EUR* and *Cleveland-EUR* cohorts

GE-PRS / *GE-Epi25*	**cases/controls upper PRS%**	**cases/controls lower PRS%**	**Sensitivity**	**Specificity**	**PPV (0.433% prevalence)**	**NPV (0.433% prevalence)**
Top 20% of distribution	887/3,652	1,369/16,783	0.393	0.821	0.009	0.997
Top 5% of distribution	305/830	1,951/19,605	0.135	0.959	0.014	0.996
Top 0.5% of distribution	54/60	2,202/20,375	0.024	0.997	0.034	0.996
GE-PRS / *GE-Cleveland*	**cases/controls upper PRS%**	**cases/controls lower PRS%**	**Sensitivity**	**Specificity**	**PPV (0.433% prevalence)**	**NPV (0.433% prevalence)**
Top 20% of distribution	35/4,070	50/16,365	0.412	0.801	0.009	0.997
Top 5% of distribution	11/1,016	74/19,419	0.129	0.950	0.011	0.996
Top 0.5% of distribution	3/100	82/20,335	0.035	0.995	0.030	0.996
FE-PRS / *FE-Epi25*	**cases/controls upper PRS%**	**cases/controls lower PRS%**	**Sensitivity**	**Specificity**	**PPV (0.299% prevalence)**	**NPV (0.299% prevalence)**
Top 20% of distribution	992/3,785	2,457/16,650	0.288	0.815	0.005	0.997
Top 5% of distribution	292/903	3,157/19,532	0.085	0.956	0.006	0.997
Top 0.5% of distribution	40/80	3,409/20,355	0.012	0.996	0.009	0.997
FE-PRS / *FE-Cleveland*	**cases/controls upper PRS%**	**cases/controls lower PRS%**	**Sensitivity**	**Specificity**	**PPV (0.299% prevalence)**	**NPV (0.299% prevalence)**
Top 20% of distribution	148/4,047	387/16,388	0.277	0.802	0.004	0.997
Top 5% of distribution	40/1,009	495/19,426	0.075	0.951	0.005	0.997
Top 0.5% of distribution	5/100	530/20,335	0.009	0.995	0.006	0.997

The positive predictive values (PPV) and negative predictive values (NPV) are calculated based on pooled point prevalence of 4.33/1000 for active generalized epilepsy, 2.99/1000 for focal epilepsy (Fiest *et al.*, 2017).

7. References

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015; 526: 68–74.

Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, van Emde Boas W, et al. Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009. Epilepsia 2010; 51: 676–685.

Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, et al. Cohort Profile: The National FINRISK Study. Int J Epidemiol 2017

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015; 4: 7.

Choi SW, Shin T, Mak H, Reilly PFO. A guide to performing Polygenic Risk Score analyses. bioRxiv 2018; 5: 11–13.

Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016; 48: 1284–1287.

Deelen P, Bonder MJ, van der Velde KJ, Westra H-J, Winder E, Hendriksen D, et al. Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. BMC Res Notes 2014; 7: 901.

Fiest KM, Sauro KM, Wiebe S, Patten SB, Kwon C-S, Dykeman J, et al. Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. Neurology 2017; 88: 296–303.

International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. Nat Commun 2018; 9: 5269.

Karlson EW, Boutin NT, Hoffnagle AG, Allen NL. Building the Partners HealthCare Biobank at Partners Personalized Medicine: Informed Consent, Return of Research Results, Recruitment Lessons and Operational Considerations. J Pers Med 2016; 6

Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 2016; 48: 1443–1448.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics (Oxford, England) 2010; 26: 2867–73.

McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature genetics 2016; 48: 1279–1283.

Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: Study design and profile. J Epidemiol 2017; 27: S2–S8.

Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008; 84: 362–369.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015; 12: e1001779.

Turchin MC, Hirschhorn JN. Gencrypt: one-way cryptographic hashes to detect overlapping individuals across samples. Bioinformatics 2012; 28: 886–888.

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011; 88: 76–82.