The road towards precision therapy for genetic epilepsies

Remi Stevelink

393

UMC Utrecht Brain Center

UMC Utrecht

Universiteit Utrecht

ISBN 978-94-6458-484-4

# The road towards precision therapy for genetic epilepsies

Remi Stevelink

# The road towards precision therapy for genetic epilepsies

**Remi Stevelink**

ISBN: 978-94-6458-484-4
Cover design: Lennart Matthijs Stokmans
Layout and design: Publiss | www.publiss.nl
Printing: Ridderprint | www.ridderprint.nl
© Copyright 2022: Remi Stevelink

Financial support for the research presented in this thesis was generously provided by the MING Fund, vrienden van WKZ.

Financial support for the printing of this thesis was provided by the MING Fund, ChipSoft, and UMC Utrecht Brain Center.

# The road towards precision therapy for genetic epilepsies

## Op weg naar precisie-behandeling van genetische epilepsie

(met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 20 oktober 2022 des middags te 4.15 uur

door

**Remi Stevelink**

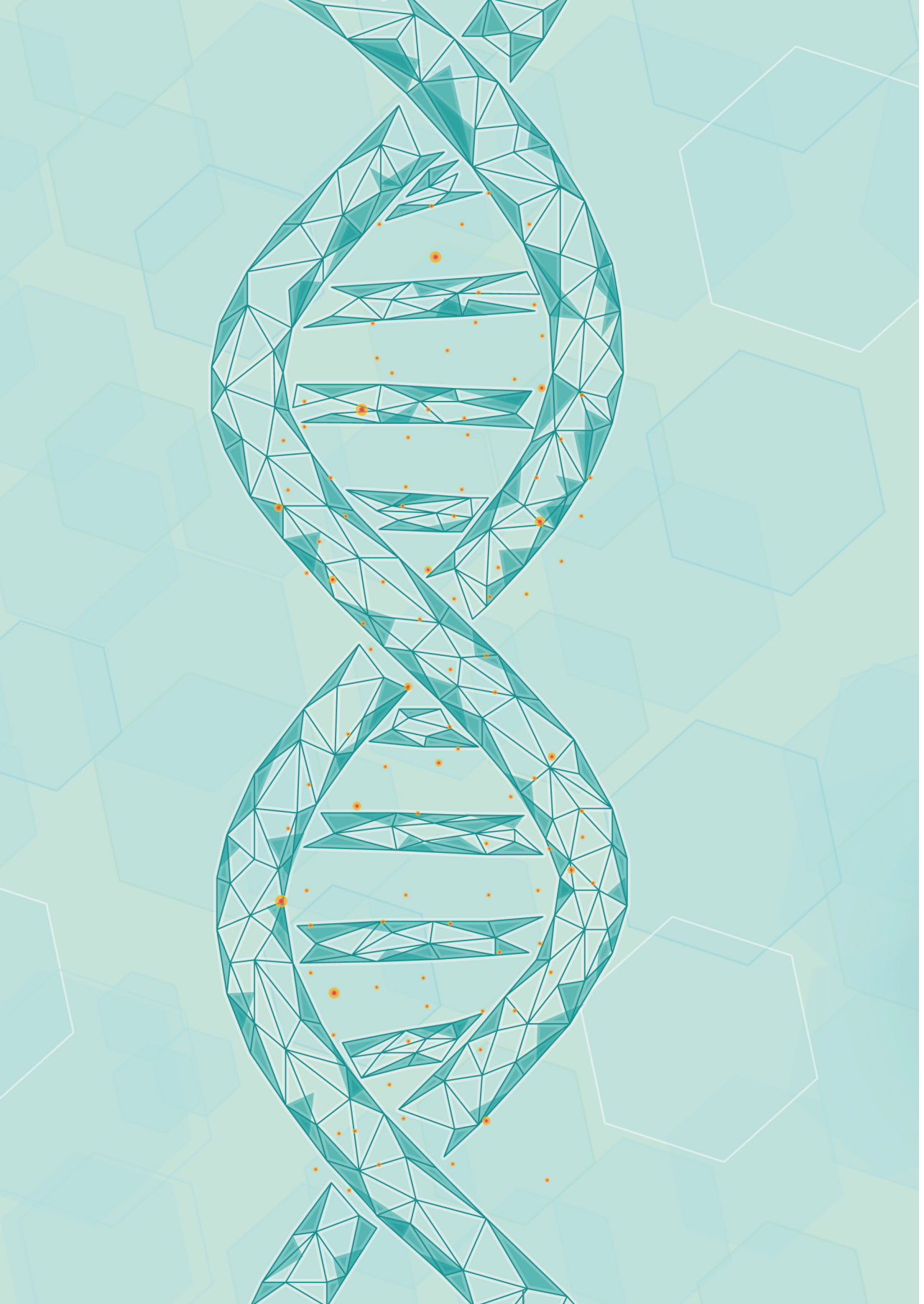
geboren op 6 oktober 1990
te Eelde

**Promotor:**
Prof. dr. K.P.J. Braun

**Copromotoren:**
Dr. B.P.C. Koeleman
Dr. F.E. Jansen

Voor Rein

# Table of contents



# CHAPTER 1

# GENERAL INTRODUCTION

Epilepsy affects around 50–70 million people worldwide, making it the most common serious neurological disorder in the world.[1,2] Although there are differences in prevalence, epilepsy occurs in people of all ages, social classes, sexes and ethnicities. Epilepsy is defined as an enduring predisposition of the brain to develop seizures and it is recognized since at least three millennia. Historically, it was often thought of as a supernatural or spiritual disorder, caused by possession by the devil.[3] We now know that epilepsy has a pathophysiological basis, caused by hyperexcitability of the brain. This hyperexcitability can affect either part of the brain, causing focal onset seizures, or both hemispheres at once, causing generalised onset seizures.[4] Seizures can have a wide variety of manifestations, depending on the location and extent of the brain involved. For example, a seizure with focal onset in the temporal lobe can cause an experience of déjà vu, while maintaining awareness. A seizure involving both hemispheres can cause contraction of muscles throughout the body due to bilateral excitation of motor neurons, and impaired awareness, since normal activity throughout the brain is disturbed. Depending on the seizure types, epilepsy can be classified as focal, generalised or a combination thereof.[5] With a combination of specific seizure types, age at onset, provoking factors, electroencephalography (EEG) findings and imaging results, it is possible to define specific electroclinical syndromes in about half of the patients. Epilepsy subtypes can be further distinguished by underlying aetiology, which is broadly grouped into structural, infectious, metabolic, immune, genetic and unknown aetiology.[5] It is increasingly recognised that understanding the aetiology of such epilepsy syndromes has important implications to improve counselling and treatment.

## Epilepsy heritability and genetics

John Russel Reynolds already recognised in 1861 that epilepsy often affects multiple family members.[7] As late as in the mid-20$^{th}$ century, such observations have led to abominable laws to forbid marriage for people with epilepsy, and even forced sterilisation of women to prevent its spread.[8] Later work by William Lennox in 1951 confirmed that epilepsy is highly heritable, by showing that many forms of epilepsy occur much more frequently in relatives of people with epilepsy, and almost always if a monozygotic twin

is affected.[9] Similar studies have later shown that almost three-quarters of people have an at least partly heritable basis of their epilepsy.[9,10] Inheritance studies showed that some forms of epilepsy have a dominant or recessive pattern of inheritance, driven by DNA changes in a single gene (called 'monogenic'), whereas the great majority of people have a more complex inheritance pattern, assumed to be caused by changes in multiple genes ('polygenic') in combination with potential environmental factors.[11]

With the advent of widespread genomic sequencing in the 1990s, it was possible to pinpoint the first epilepsy genes causing monogenic familial epilepsy.[9] Further advances in accuracy, speed and affordability of sequencing technology have resulted in the discovery of hundreds of epilepsy genes. Furthermore, we now know that many of the most severe forms of genetic epilepsy (so called 'developmental and epileptic encephalopathies') are not mendelian inherited, but occur *de novo* in the gametes of the parent, or in the developing embryo.[12] Another surprise was the finding of genetic causes of not only generalised but also some forms of focal epilepsy. Although monogenic types of epilepsy are individually very rare, collectively they now add up to a considerable proportion of people with epilepsy. A further proportion of epilepsy genetics can be explained by copy number variants (CNV), i.e. deletions or duplications of large chunks of the genome. Such CNVs can increase epilepsy risk, or even be sufficient to cause epilepsy by themselves.[13]

Despite enormous advances in epilepsy gene discovery, the large majority of common epilepsy heritability remains unexplained. In particular, the group of genetic generalised epilepsies (GGE) – also known as idiopathic generalised epilepsy – is long known to be highly heritable. The most common GGE syndrome is juvenile myoclonic epilepsy (JME), which is characterised by an adolescent age at onset, myoclonic seizures provoked by sleep deprivation, and 4 to 6 Hz polyspikes and waves on EEG, without abnormalities on structural neuroimaging.[14–16] Twin based studies have estimated heritability of GGE to be around 65-76%.[17] However, discovery of GGE genes has remained elusive. Although various monogenic GGE candidate genes have been identified,[18–20] such genes can only explain a small proportion of cases. Instead, it is increasingly understood that the majority of GGE heritability is polygenic in nature, where multiple common genetic variants together explain the heritable risk of epilepsy.

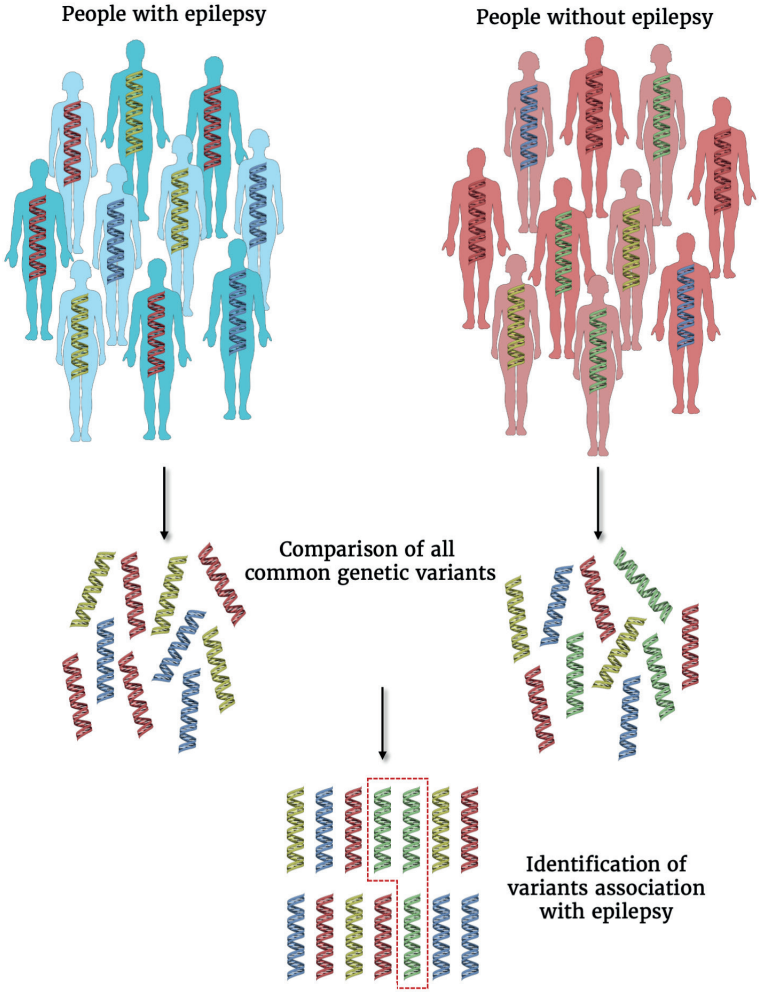



**Figure 1: The concept of a genome-wide association study (GWAS).** The aim of a GWAS is to identify the genetic variants that underlie a disease, in this case epilepsy. This is performed by comparing all common variants in the genome (generally several millions) between large groups (generally several thousands) of people with epilepsy with people without epilepsy, while correcting for potential confounders such as population stratification. Since around 1 million independent variants are assessed, a stringent significance threshold ($p=5*10^{-8}$) is set to correct for multiple testing. After identifying epilepsy risk variants, downstream analyses can elucidate involved genes, pathophysiological mechanisms and potential drug targets. Figure created using images from smart.servier.com.

Discovering common genetic variants through genetic association studies requires large sample sizes, since these variants individually only have a small effect on epilepsy risk. Common genetic variants are usually identified through

genome-wide association studies (GWAS), which utilize large sample sizes to compare the frequency of all common genetic variants in the population. Typically, thousands of subjects are required to identify risk variants at nominal significance, after correction for around 1 million independent tested variants. The first epilepsy GWAS meta-analysis identified three risk variants by increasing sample sizes through large-scale collaborations,[21] however, these still only explain a negligible fraction of epilepsy risk. Further increases in sample size in GWAS of other diseases have already shown that genetic data from tens of thousands, or even millions of individuals, can drastically increase the amount of risk loci and explained heritability.[22] Such findings can be used for clinical prediction, increased understanding of disease pathophysiology and lead to novel therapeutic approaches.[22]

## Epilepsy treatment

The mainstay of epilepsy treatment consists of anti-seizure medication (ASM). Such drugs generally work by decreasing brain excitability, which could also lead to adverse effects.[24] In a large study, 88% of people reported one or more adverse effects, which negatively impacted on quality of life.[23] When ASM therapy fails, surgical resection of a lesion can be a curative option for a subset of people with focal epilepsy, however, this is not an option for generalised epilepsy. Further non-pharmacological treatment options include vagal nerve stimulation and the ketogenic diet, however, such treatment options also suffer from limitations in applicability and efficacy.[26] Although the majority of people with epilepsy can be effectively treated with ASMs, around a third continue to have seizures. This proportion of drug resistance remained stable over the last decades, despite the availability of numerous new drugs.[27] One likely reason is that many of the current ASMs were found serendipitously, and many of the ASMs work by similar mechanisms. Potentially, new therapeutic targeted approaches, informed by underlying pathophysiology, could increase the proportion of people becoming seizure free.

## Epilepsy prediction

Improved prediction could further improve and personalise treatment, within the arsenal of current treatment options. Individualized prediction is

an important part of modern precision medicine approaches. Prediction can be broadly categorised into diagnostic and prognostic prediction. Diagnostic prediction includes an earlier diagnosis of epilepsy or an earlier or more specific diagnosis of an epilepsy subtype. For example, after a first event – possibly a seizure – it is often difficult to know if the patient will eventually be diagnosed with epilepsy or not. Early and accurate diagnostic prediction can accelerate treatment initiation in individuals highly suspected of having epilepsy, and prevent unnecessary medication in those not at risk. A prediction model based on information known after a first event could aid an earlier diagnosis and treatment.[27,28] Genetic diagnostics have the potential to further expedite predictions. Since genetic variants generally remain stable throughout life, it is possible to use genetics to predict epilepsy risk even before someone has a first seizure. Theoretically, such a genetic predisposition for epilepsy could be detected even by newborn screening technologies. This could be of especial relevance for people with a familial burden of epilepsy.

Improved prediction of prognosis will aid personalised treatment and counselling of people with epilepsy. Currently, the standard approach of epilepsy treatment is trial-and-error.[23] People are prescribed a drug that is selected based on efficacy in broad populations. When the drug fails to control seizures, a second or more consecutive drugs are attempted. This process often takes years, during which a patient might experience many debilitating seizures and adverse drug reactions, and it is currently impossible to predict the chances of any individual to become seizure-free with treatment. Improved prediction of drug resistance could expedite triage of people at risk to specialised care and improve treatment outcome.[27,28] For people who have managed to become free of seizures while on ASM, it is often a difficult dilemma whether to continue or withdraw ASM medication. Continuing treatment could maximise chances of remaining seizure free and eligibility for a driving license, however, it comes at the cost of taking daily medication with potential adverse effects.[30,31] Predicting risk of seizure relapse after withdrawal could guide who should continue treatment and who could safely withdraw treatment, thereby potentially improving quality of life.[14–16] Such predictions are already available based on broad cohorts of people with epilepsy,[32] however, predictions could potentially be improved by basing them on people with the same specific subtypes of epilepsy. In

addition to clinical predictors, genetics could also aid prognosis prediction, since people with the same genetic aetiology often have similar treatment outcomes. This is currently of most relevance for monogenic epilepsies. For example, sodium channel blocking drugs like phenytoin can be very effective to treat epilepsy due to gain-of-function variants in the sodium channel gene SCN8A, whereas they could worsen seizures in people with epilepsy due to loss-of-function SCN1A variants.

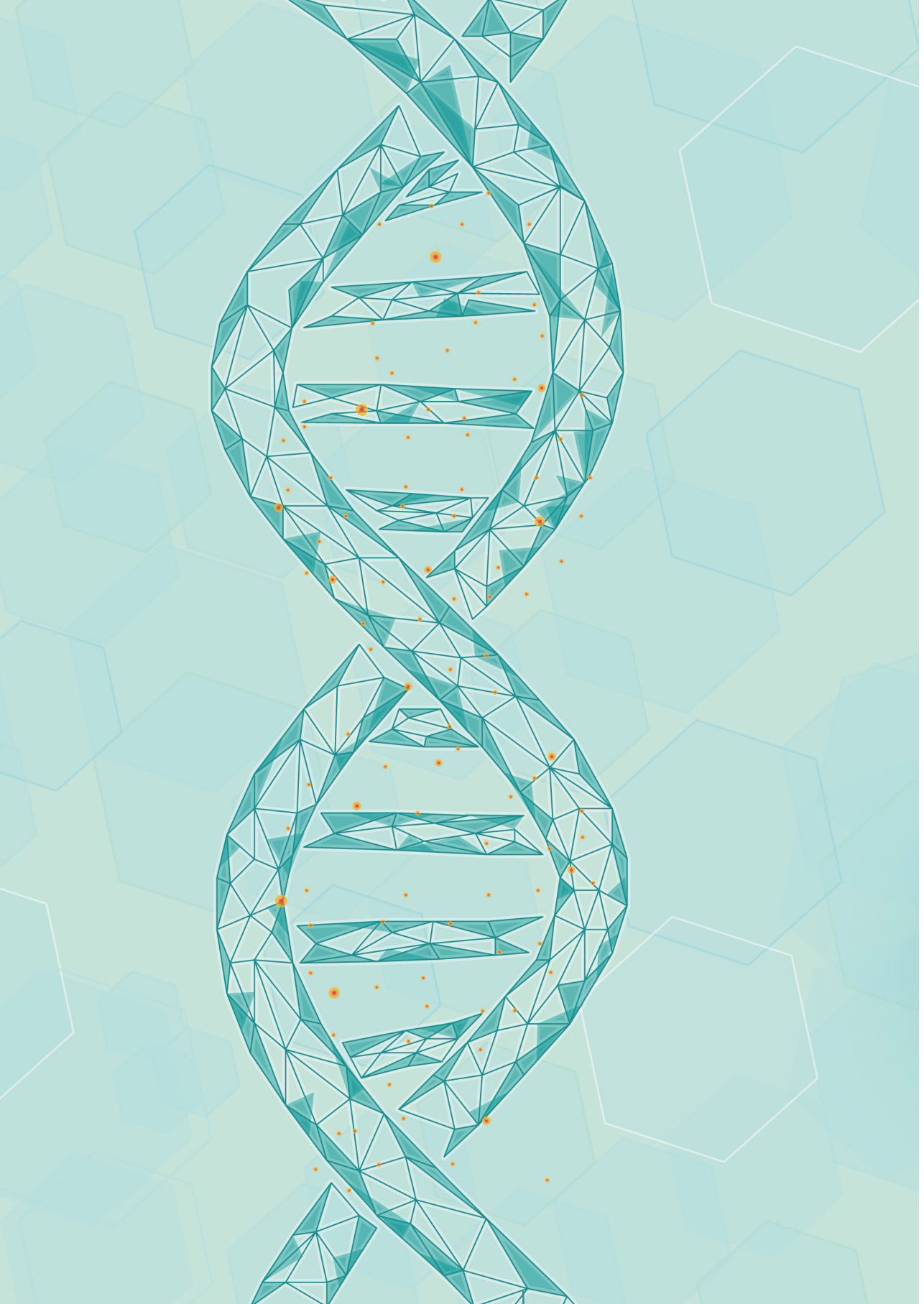
## Outline and aims of this thesis

The overarching aim of this thesis is to increase understanding of genetic risk factors for epilepsy, and to use this knowledge to improve and personalize epilepsy treatment. We employed a broad arsenal of research techniques that we hope will aid to pave the road towards precision therapy for common genetic epilepsies. In **chapter 2** we describe a large-scale GWAS of common epilepsies, where we identified 16 epilepsy risk loci to improve our understanding of epilepsy heritability and pathophysiology. We followed-up on a lead from our GWAS and assessed a potential genetic association between GGE and altered vitamin-B6 metabolism in **chapter 3**. In **chapter 4**, we assessed whether there is a genetic relation between epilepsy and background EEG measures, which could potentially aid epilepsy diagnosis. The aim of **chapter 5** was to assess how GWAS data could be used to find effective epilepsy drugs and guide drug repurposing. **Chapter 6** describes a GWAS with an almost doubled sample size compared to chapter 2, which further improves our understanding of epilepsy. We used this increased sample size in **chapter 7** to assess whether differences on brain MRI scans in people with epilepsy reflect the cause or consequence of epilepsy or its treatment. In **chapter 8**, we investigated whether common genetic variants identified in GWAS, combined in polygenic risk scores, could improve diagnosis and classification of common epilepsies. **Chapter 9** describes how knowledge about monogenic focal epilepsies could guide which patients might benefit from epilepsy surgery. In **chapter 10**, we assessed how often people with JME are drug-resistant and how many people experience a relapse after ASM withdrawal. We built upon this in **chapter 11**, by creating and validating individualised prediction models of treatment outcomes, which can be of use to personalise treatment and counselling of people with JME.

# References


1. Thurman DJ, Beghi E, Begley CE, *et al.* Standards for epidemiologic studies and surveillance of epilepsy. *Epilepsia* 2011; **52 Suppl 7**: 2–26.
2. Reynolds EH. The ILAE/IBE/WHO epilepsy global campaign history. International League Against Epilepsy. International Bureau for Epilepsy. *Epilepsia* 2002; **43 Suppl 6**: 9–11.
3. Pierce JMS. A DISEASE ONCE SACRED. A HISTORY OF THE MEDICAL UNDERSTANDING OF EPILEPSY. *Brain* 2002; **125**: 441–2.
4. Fisher RS, Cross JH, French JA, *et al.* Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 2017; **58**: 522–30.
5. Scheffer IE, Berkovic S, Capovilla G, *et al.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 2017; **58**: 512–21.
6. Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. *Epilepsia* 1989; **30**: 389–99.
7. Reynolds SJR. Epilepsy-- its symptoms, treatment, and relation to other chronic convulsive diseases. Churchill, 1861.
8. Kaculini CM, Tate-Looney AJ, Seifi A. The History of Epilepsy: From Ancient Mystery to Modern Misconception. *Cureus* 2021; **13**: e13953.
9. Lennox WG. The heredity of epilepsy as told by relatives and twins. *J Am Med Assoc* 1951; **146**: 529–36.
10. Thomas RH, Berkovic SF. The hidden genetics of epilepsy-a clinically important new paradigm. *Nat Rev Neurol* 2014; **10**: 283–92.
11. Tsuboi T, Christian W. On the genetics of the primary generalized epilepsy with sporadic myoclonias of impulsive petit mal type. A clinical and electroencephalographic study of 399 probands. *Humangenetik* 1973; **19**: 155–82.
12. Helbig I, Tayoun AAN. Understanding Genotypes and Phenotypes in Epileptic Encephalopathies. *Mol Syndromol* 2016; **7**: 172–81.
13. Mefford HC. Copy Number Matters in Epilepsy. *Epilepsy Curr* 2015; **15**: 180–2.
14. Kjeldsen MJ, Corey LA, Christensen K, Friis ML. Epileptic seizures and syndromes in twins: the importance of genetic factors. *Epilepsy Res* 2003; **55**: 137–46.
15. Berkovic SF, Howell RA, Hay DA, Hopper JL. Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann Neurol* 1998; **43**: 435–45.
16. Corey LA, Pellock JM, Kjeldsen MJ, Nakken KO. Importance of genetic factors in the occurrence of epilepsy syndrome type: a twin study. *Epilepsy Res* 2011; **97**: 103–11.
17. Delgado-Escueta AV, Koeleman BPC, Bailey JN, Medina MT, Durón RM. The quest for juvenile myoclonic epilepsy genes. *Epilepsy Behav* 2013; **28 Suppl 1**: S52-7.
18. EPICURE Consortium, EMINet Consortium, Steffens M, *et al.* Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. *Hum Mol Genet* 2012; **21**: 5359–72.

19. International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address: epilepsy-austin@unimelb.edu.au. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol* 2014; **13**: 893–903.
20. Kasperaviciūte D, Catarino CB, Heinzen EL, *et al.* Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. *Brain* 2010; **133**: 2136–47.
21. Wightman DP, Jansen IE, Savage JE, *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet* 2021; **53**: 1276–82.
22. Uffelmann E, Huang QQ, Munung NS, *et al.* Genome-wide association studies. *Nature Reviews Methods Primers* 2021; **1**: 1–21.
23. Perucca P, Carter J, Vahle V, Gilliam FG. Adverse antiepileptic drug effects: toward a clinically and neurobiologically relevant taxonomy. *Neurology* 2009; **72**: 1223–9.
24. Devinsky O, Vezzani A, O'Brien TJ, *et al.* Epilepsy. *Nat Rev Dis Primers* 2018; **4**: 18024.
25. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment Outcomes in Patients With Newly Diagnosed Epilepsy Treated With Established and New Antiepileptic Drugs: A 30-Year Longitudinal Cohort Study. *JAMA Neurol* 2018; **75**: 279–86.
26. van Diessen E, Lamberink HJ, Otte WM, *et al.* A Prediction Model to Determine Childhood Epilepsy After 1 or More Paroxysmal Events. *Pediatrics* 2018; **142**. DOI:10.1542/peds.2018-0931.
27. Chen Z, Rollo B, Antonic-Baker A, *et al.* New era of personalised epilepsy management. *BMJ* 2020; **371**: m3658.
28. Szaflarski JP, Rackley AY, Lindsell CJ, Szaflarski M, Yates SL. Seizure control in patients with epilepsy: the physician vs. medication factors. *BMC Health Serv Res* 2008; **8**: 264.
29. Sillanpää M, Haataja L, Shinnar S. Perceived impact of childhood-onset epilepsy on quality of life as an adult. *Epilepsia* 2004; **45**: 971–7.
30. Lamberink HJ, Boshuisen K, Otte WM, Geleijns K, Braun KPJ, TimeToStop Study Group. Individualized prediction of seizure relapse and outcomes following antiepileptic drug withdrawal after pediatric epilepsy surgery. *Epilepsia* 2018; **59**: e28–33.
31. Lamberink HJ, Otte WM, Geerts AT, *et al.* Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. *Lancet Neurol* 2017; **16**: 523–31.



## CHAPTER 2

# GENOME-WIDE MEGA-ANALYSIS IDENTIFIES 16 LOCI AND HIGHLIGHTS DIVERSE BIOLOGICAL MECHANISMS IN THE COMMON EPILEPSIES

The International League Against Epilepsy Consortium on Complex Epilepsies
Contribution: Co-lead author and analist

*Nature Communications*. 2018; Dec 10;9(1):5269.

## Abstract

The epilepsies affect around 65 million people worldwide and have a substantial missing heritability component. We report a genome-wide mega-analysis involving 15,212 individuals with epilepsy and 29,677 controls, which reveals 16 genome-wide significant loci, of which 11 are novel. Using various prioritization criteria, we pinpoint the 21 most likely epilepsy genes at these loci, with the majority in genetic generalized epilepsies. These genes have diverse biological functions, including coding for ion-channel subunits, transcription factors and a vitamin-B6 metabolism enzyme. Converging evidence shows that the common variants associated with epilepsy play a role in epigenetic regulation of gene expression in the brain. The results show an enrichment for monogenic epilepsy genes as well as known targets of antiepileptic drugs. Using SNP-based heritability analyses we disentangle both the unique and overlapping genetic basis to seven different epilepsy subtypes. Together, these findings provide leads for epilepsy therapies based on underlying pathophysiology.

# Introduction

The epilepsies are a group of brain disorders characterized by recurrent unprovoked seizures affecting up to 65 million people worldwide[1]. There are many different types of epilepsy, and its classification has recently evolved, driven by advances in clinical phenotyping, imaging, and genetics[2]. Since the identification of *CHRNA4* as a cause of autosomal dominant nocturnal frontal lobe epilepsy[3], genes underlying many different rare monogenic forms of epilepsy have been characterized, and discovery in this area has accelerated with the application of next generation sequencing[4]. This is particularly true of the relatively rare but devastating infantile group of epileptic encephalopathies, which are now emerging as a genetically heterogeneous group of largely de novo dominant disorders[5]. In contrast, single gene causes of the more common forms of epilepsy appear to be relatively rare. The common forms broadly comprise generalized and focal epilepsies, with the former having the highest heritability, yet the lesser yield in single gene discovery[6]. These common forms are likely multifactorial, with a significant and complex genetic architecture[7,8,9].

Consistent with the experience from many other disease fields, early attempts to disentangle the genetic architecture of the more common, sporadic forms of epilepsy were limited by study power and scope[10,11,12,13,14]. In 2011, the International League Against Epilepsy (ILAE) launched the Consortium on Complex Epilepsies, to facilitate meta-analysis in epilepsy genomics. In 2014, the first such meta-analysis was reported comprising 8696 cases and 26,157 controls. This led to the identification of 2q24.3, 4p15.1, and 2p16.1 as epilepsy loci[15].

Here we present an expanded analysis involving 15,212 cases and 29,677 controls, which leads to identification of 16 genome-wide significant loci. Importantly, 11 of these loci are associated with the genetic generalized epilepsies; the group of epilepsies where despite having the highest heritability we have made the least genetic progress to date. We show that the genes associated with each locus are biologically plausible candidates, despite having diverse functions, particularly as there is a significant enrichment for known monogenic epilepsy genes and antiepileptic drug targets.

# Results

## Study overview

We performed a genome-wide mega-analysis on the ILAE Consortium cohort now comprising 15,212 epilepsy cases, stratified into 3 broad and 7 subtypes of epilepsy, and 29,677 control subjects (Supplementary Table 1). The current study includes a further 6516 cases and 3460 controls in addition to the 8696 cases and 26,157 controls from our previously published analysis[15]. Thus, this mega-analysis is not a formal replication of our previously published meta-analysis. We do not attempt any formal replication of novel association signals detected in this analysis. Furthermore, 531 cases of Asian descent, and 147 cases of African descent were included through a meta-analysis. However, we refer to our GWAS as a mega-analysis as the vast majority of our samples (96%) were analyzed under that framework.

At the broadest level, cases were classified as (a) focal epilepsy where seizures arise in a restricted part of the brain, a form traditionally not regarded as genetic although a number of genes for monogenic forms have been identified; (b) genetic generalized epilepsy where seizures arise in bilateral networks and evidence for a genetic component is very strong, yet genes have been hard to identify, and (c) unclassified epilepsy[2,16].

Subjects were assigned to three broad ancestry groups (Caucasian, Asian and African-American) according to results of genotype-based principal component analysis (Supplementary Fig. 1). Linear-mixed model analyses were performed stratified by ethnicity and epilepsy subtype or syndrome, after which trans-ethnic meta-analyses were undertaken.

## Genome-wide associations

Our analysis of all epilepsy cases combined revealed one novel genome-wide significant locus at 16q12.1 and reinforced two previous associations at 2p16.1 and 2q24.3 (Fig. 1 and Supplementary Fig. 2)[15]. When conditioning on the top SNP within the 2q24.3 locus, we demonstrate the existence of a second, independent signal within that locus (Supplementary Fig. 3). This locus was also significantly associated with focal epilepsy. Our analysis of

genetic generalized epilepsy uncovered 11 genome-wide significant loci, of which seven are novel (Fig. 2).

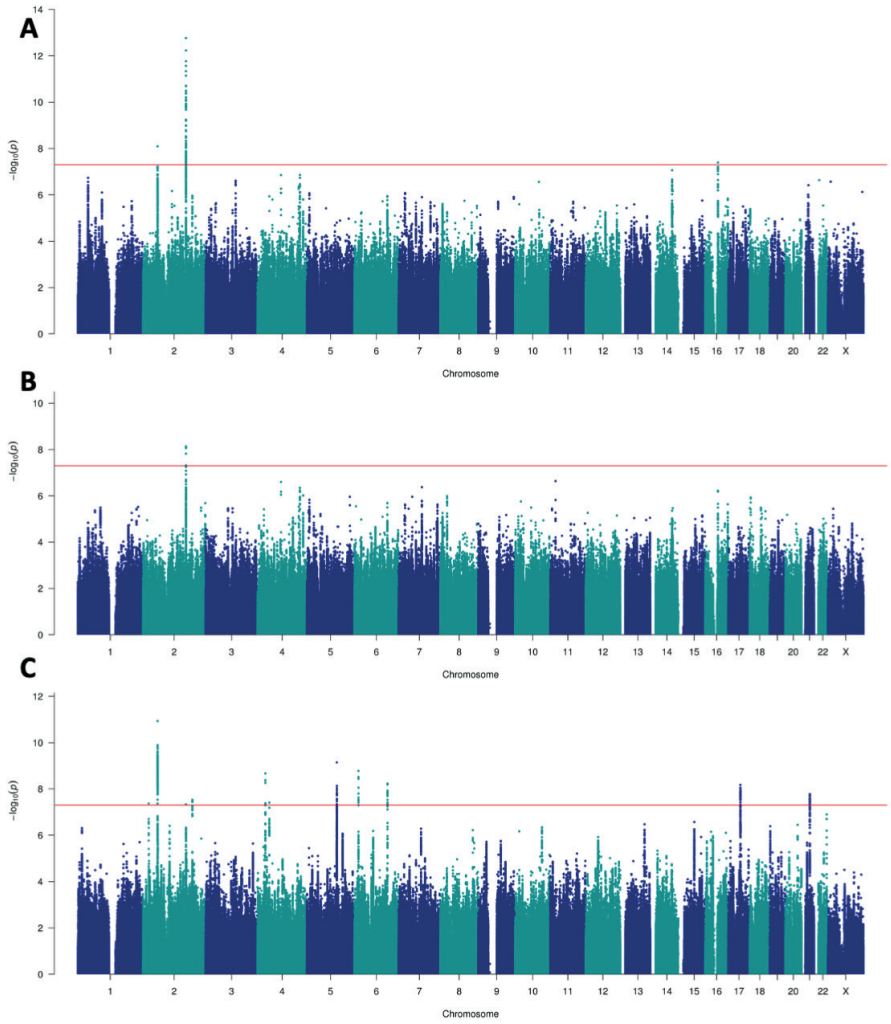



**Figure 1: Manhattan plots for epilepsy genome-wide association analyses**. Genome-wide association analyses of **a** all epilepsy, **b** focal epilepsy, and **c** genetic generalized epilepsy. Negative log10-transformed P-values (Y-axis) are plotted against chromosomal position (x-axis). P-values were calculated with METAL using fixed-effects trans-ethnic meta-analyses. The red line represents the genome-wide significance threshold ($p < 5 \times 10-8$). Previously known loci are indicated in black; novel loci in red. The names above each locus represent the prioritized gene in the locus (see Fig. 2) or the name of the locus itself in case of multiple prioritized genes in the locus

| Phenotype | Locus | Novel/replication | Lead SNP | MAF | *Z*-score | *P*-value | Gene | Total score | Biological gene criteria | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | TWAS | eQTL | Brain exp | Missense | PPI | KO mouse | AED target | Monogenic |
| All epilepsy | 2p16.1 | Replication | rs4671319 (G) | 0.44 | 5.77 | 8.1E–09 | FANCL | 2 | ■ | | ■ | | | | | |
| | | | | | | | BCL11A | 2 | | | ■ | | | ■ | | |
| | 2q24.3 | Replication | rs6432877 (G) | 0.26 | 7.37 | 1.7E–13 | SCN3A | 3 | | | ■ | | ■ | ■ | ■ | |
| | | | | | | | SCN2A | 3 | | | ■ | | ■ | ■ | ■ | ■ |
| | | | | | | | TTC21B | 3 | | ■ | | ■ | | ■ | | |
| | | | | | | | SCN1A | 3 | | | ■ | ■ | | ■ | ■ | ■ |
| | 16q12.1 | Novel | rs4638568 (A) | 0.06 | –5.49 | 4.0E–08 | HEATR3 | 2 | | ■ | | | ■ | | | |
| | | | | | | | BRD7 | 2 | | | | | | | | |
| Focal epilepsy | 2q24.3 | Replication | rs2212656 (A) | 0.26 | 5.78 | 7.3E–09 | SCN3A | 3 | | | ■ | | ■ | ■ | ■ | |
| | | | | | | | SCN2A | 3 | | | ■ | | ■ | ■ | ■ | ■ |
| | | | | | | | TTC21B | 3 | | ■ | | ■ | | ■ | | |
| | | | | | | | SCN1A | 3 | | | ■ | ■ | | ■ | ■ | ■ |
| Generalized epilepsy | 2p24.1 | Novel | rs4665630 (C) | 0.13 | 5.48 | 4.3E–08 | None | | | | | | | | | |
| | 2p16.1 | Replication | rs1402398 (G) | 0.36 | 6.79 | 1.2E–11 | FANCL | 2 | ■ | | ■ | | | | | |
| | | | | | | | BCL11A | 2 | | | ■ | | | ■ | | |
| | 2q24.3 | Replication | rs11890028 (G) | 0.27 | –5.46 | 4.7E–08 | SCN3A | 3 | | | ■ | | ■ | ■ | ■ | |
| | | | | | | | SCN2A | 3 | | | ■ | | ■ | ■ | ■ | ■ |
| | | | | | | | TTC21B | 3 | | ■ | | ■ | | ■ | | |
| | | | | | | | SCN1A | 3 | | | ■ | ■ | | ■ | ■ | ■ |
| | 2q32.3 | Novel | rs887696 (C) | 0.34 | 5.54 | 3.0E–08 | STAT4 | 2 | | | ■ | | ■ | | | |
| | 4p15.1 | Replication | rs1044352 (T) | 0.42 | 5.98 | 2.2E-09 | PCDH7 | 2 | | | ■ | | | | | |
| | 4p12 | Novel | rs11943905 (T) | 0.27 | 5.50 | 3.9E–08 | GABRA2 | 4 | ■ | | ■ | | ■ | ■ | ■ | |
| | 5q22.3 | Novel | rs4596374 (C) | 0.45 | 6.16 | 7.2E–10 | KCNN2 | 2 | | | ■ | | | ■ | | |
| | 6p22.3 | Novel | rs68082256 (A) | 0.20 | –6.02 | 1.7E–09 | ATXN1 | 2 | | | ■ | | | ■ | | |
| | 6q22.33 | Novel | rs13200150 (G) | 0.30 | –5.82 | 5.9E–09 | None | | | | | | | | | |
| | 17q21.32 | Replication | rs4794333 (C) | 0.38 | –5.80 | 6.8E–09 | PNPO | 3 | ■ | | ■ | | ■ | | | ■ |
| | 21q22.11 | Novel | rs2833098 (G) | 0.38 | –5.64 | 1.7E–08 | GRIK1 | 2 | | | ■ | | | ■ | ■ | |
| JME | 16p11.2 | Novel | rs1046276 (T) | 0.34 | 6.67 | 2.5E–11 | STX1B | 4 | | ■ | ■ | | ■ | ■ | | ■ |
| CAE | 2p16.1 | Replication | rs12185644 (C) | 0.29 | 6.24 | 4.5E–10 | FANCL | 2 | ■ | | ■ | | | | | |
| | | | | | | | BCL11A | 2 | | | ■ | | | ■ | | |
| | 2q22.3 | Replication | rs13020210 (G) | 0.20 | –5.58 | 2.4E–08 | ZEB2 | 2 | | | ■ | | | ■ | | ■ |
| Focal HS | 3q25.31 | Novel | rs1991545 (A) | 0.04 | 6.78 | 1.3E–11 | C3orf33 | 2 | | | ■ | | ■ | | | |
| | | | | | | | SLC33A1 | 2 | | | ■ | | | ■ | | |
| | | | | | | | KCNAB1 | 2 | | | ■ | | | ■ | | |
| | 6q22.31 | Novel | rs1318322 (G) | 0.14 | 5.80 | 6.7E–09 | GJA1 | 2 | | | ■ | | | ■ | | |

**Figure 2: Genome-wide significant loci of all analyses and prioritized biological epilepsy genes.** Genes were prioritized based on 6 criteria and scored based on the number of criteria met per gene (filled red boxes). The highest scoring gene, or multiple if they have the same score, in each locus is reported as 'prioritized biological epilepsy gene(s)'. Similar to previous studies17,18, we used a minimum score of 2 to define these genes and we noted 'none' if no gene in the locus reached this score. Filled blue boxes indicate overlap with known targets of anti-epileptic drugs and established monogenic epilepsy genes. The lead SNP is defined as the SNP with the lowest P-value in the locus and the minor allele is displayed in brackets. P-values and Z-scores for All epilepsy, Focal epilepsy and Generalized epilepsy were calculated with fixed-effects trans-ethnic meta-analyses. P-values and Z-scores for JME, CAE, and Focal HS were calculated with BOLT-LMM. MAF minor allele frequency in the Human Reference Consortium reference panel. The direction of the Z-score is signed with respect to the minor allele. TWAS: significant TWAS association (based on data from the CommonMind Consortium), eQTL: significant eQTL within locus (based on data from the ROS/MAP projects), Brain exp: the gene is preferentially expressed in the brain, Missense: epilepsy GWAS missense variant in locus, PPI: gene prioritized by protein-protein interaction, KO mouse: relevant knockout mouse phenotype.

Considering that focal and generalized epilepsy are clinically broad and heterogeneous classifications, we next assessed whether loci are specifically associated with any of the seven most common focal epilepsy phenotypes and genetic generalized epilepsy syndromes (Supplementary Fig. 4 and 5). We found a novel genome-wide significant association with

juvenile myoclonic epilepsy (JME) and two novel loci associated with focal epilepsy with hippocampal sclerosis. Moreover, we found two genome-wide significant associations with childhood absence epilepsy (CAE) in loci that were previously associated with absence epilepsy and generalized epilepsy[12]. We did not find any significant loci associated with generalized epilepsy with tonic-clonic seizures (GTCS) alone, juvenile absence epilepsy (JAE), lesion-negative or lesional focal epilepsy (other than hippocampal sclerosis). Further analysis of the association signals for each locus in the different syndromes suggested that some signals display specificity for a single subtype, while others show evidence for pleiotropy (Supplementary Fig. 6). However, the relatively small sample sizes of these phenotype subsets warrant caution for over-interpretation.

In total, we found 11 novel genome-wide significant loci associated with the epilepsies and we replicated the association of five previous known loci[12,15] (Supplementary Fig. 7). Two previous reports of association did not reach our threshold for significance. This included a locus (rs2292096; 1q32.1) for focal epilepsy detected in an Asian population 14 ($p = 0.057$ in our trans-ethnic fixed-effects meta-analysis), and rs12059546 (1q43) detected previously for JME[12] ($p = 7.4 \times 10^{-5}$ in our Caucasian-only BOLT-LMM analysis).

## Gene mapping and biological prioritization

The genome-wide significant loci from all analyses were mapped to a total of 146 genes (Supplementary Data 1) using a combination of positional mapping (±250 kb from locus) and significant distal 3D chromatin interactions of the locus with a gene promoter (FDR $< 10^{-6}$). Considering that most loci encompass several genes, we devised criteria to systematically prioritize the most likely candidate genes per locus based on established bioinformatics methods and resources. This biological prioritization was based on six criteria (Fig. 2), similar to previous studies[17,18]. Each gene was given a score based on the number of criteria that were met (range 0–6). The gene(s) with the highest score in each locus, with a minimum of 2, were defined as biological epilepsy risk genes. We validated this approach using established epilepsy genes within our data (Supplementary Table 2). Using this approach, we were able to refine these loci to the 21 most likely biological epilepsy genes (Fig. 2).

These prioritized genes include seven ion-channel genes (*SCN1A*, *SCN2A*, *SCN3A*, *GABRA2*, *KCNN2*, *KCNAB1*, and *GRIK1*), three transcription factors (*ZEB2*, *STAT4* and *BCL11A*), the histone modification gene *BRD7*, the synaptic transmission gene *STX1B* and the pyridoxine metabolism gene *PNPO*. Notably, a conditional transcriptome-wide association study (TWAS) analysis suggests that the signal for genetic generalized epilepsy at 17q21.32, which was also observed in an earlier study[12], seems driven by regulation of expression of *PNPO* (Supplementary Fig. 8). This suggests that the biology behind pyridoxine (vitamin-B6)-responsive epilepsy[19,20] could be relevant to common genetic generalized epilepsies. Biological prioritization implicates *SCN1A*, *SCN2A*, *SCN3A*, and *TTC21B* as the most likely genes underlying the signal at 2q24.3 for all epilepsy, focal epilepsy and genetic generalized epilepsy. Pathogenic variants in the sodium channels *SCN1A*, *SCN2A* and *SCN3A* are associated with various epilepsy syndromes[16] and mutations in *TTC21B* impair forebrain development[21,22]. Our analyses implicate *STX1B* as a potential gene underlying the association of JME at the 16p11.2 locus and the top variant in the locus is an eQTL that strongly correlates with expression of *STX1B* in the dorsolateral prefrontal cortex (Spearman's correlation: Rho = 0.33, $p = 3 \times 10^{-14}$)[23]. Interestingly, for one of the prioritized genes in genetic generalized epilepsy, *PCDH7*, an eQTL was recently detected in epileptic hippocampal tissue[24]. Prioritized genes associated with focal epilepsy with hippocampal sclerosis include the gap-junction gene *GJA1*.

In addition we identified eight genes from Fig. 2 (*BCL11A*, *GJA1*, *ATXN1*, *GABRA2*, *KCNAB1*, *SCN3A*, *PCDH7*, *STAT4*) with evidence of co-expression in at least two independent brain expression resources, using a brain gene co-expression analysis with brain-coX[25]. These eight candidates are embedded in several established epilepsy gene co-expression modules (Supplementary Fig. 9; Supplementary Table 9).

## SNP annotation and tissue-specific partitioned heritability

We functionally annotated all 492 genome-wide significant SNPs from all phenotypes (Fig. 3a–c) and found that most SNPs were either intergenic

(29%) or intronic (46%); 78% were in open chromatin regions (as indicated by a minimum chromatin state of 1–7[26,27], and 50% of SNPs showed some

evidence for affecting gene transcription (RegulomeDB score ≤6[28]). Four were coding SNPs of which two were missense variants.

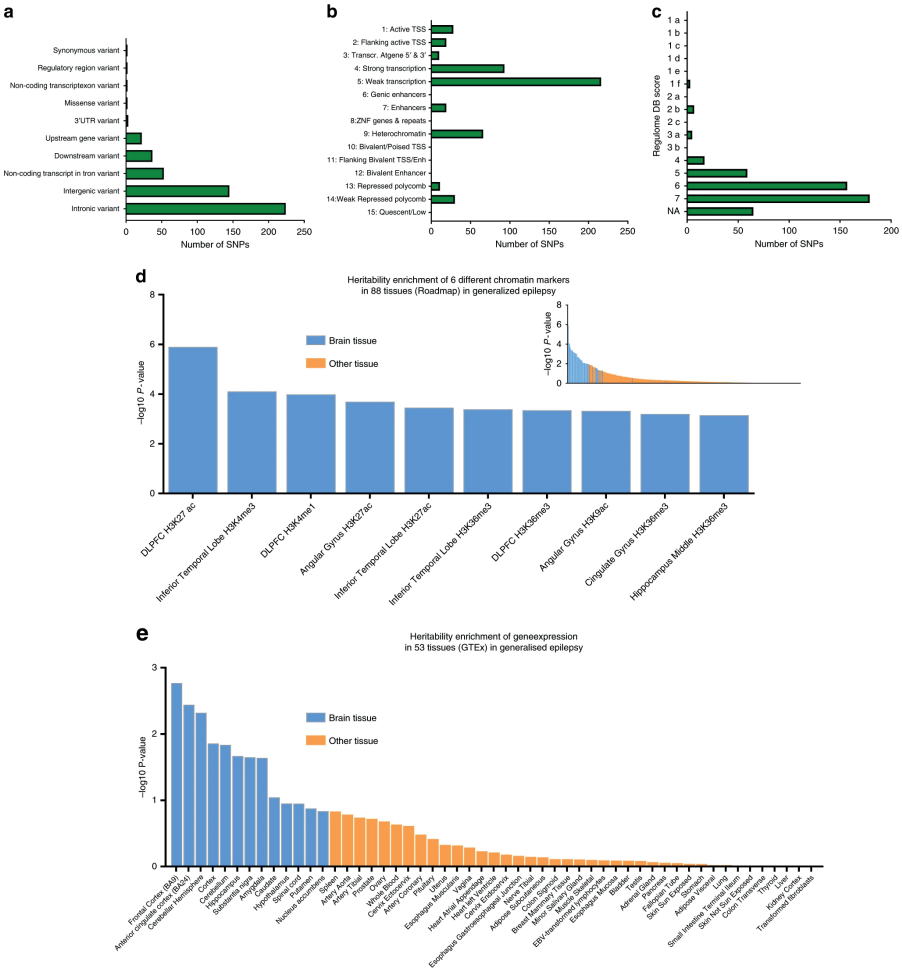



**Figure 3: Functional annotation and heritability enrichment of epilepsy GWAS results. a** functional categories of all genome-wide significant SNPs in all phenotypes. **b** Minimum (most active) chromatin state across 127 tissues for all genome-wide significant SNP in all phenotypes; TSS - transcription start site. **c** The RegulomeDB score for all genome-wide significant SNPs in all phenotypes, where 7 represents no evidence for affecting regulation and lower scores represent increasing evidence; NA - the variant does not exist in RegulomeDB. **d** Heritability enrichment for genetic generalized epilepsy with 6 different chromatin markers in 88 tissues, calculated with stratified LD-score regression using data from the Roadmap Epigenomics Project. The main bar chart represent the 10 tissues with the strongest heritability enrichment and the inset shows the full distribution of all chromatin markers in all tissues. **e** Heritability enrichment of genes expressed in 53 tissues, calculated with stratified LD-score regression using data from the gene-tissue expression (GTEx) Consortium.

To gain further biological insight into our results, we next used a partitioned heritability method[29] to assess whether our genome-wide significant signals contained SNPs associated with enhanced transcription in any of 88 tissues. We found significant enrichment of H3K4me1 markers in all epilepsy (stratified LD-score regression; $p = 4 \times 10^{-5}$) and H3K27ac markers in genetic generalized epilepsy (stratified LD-score regression; $p = 1.3 \times 10^{-6}$), specifically in the dorsolateral prefrontal cortex. Moreover, the distribution of heritability enrichment P-values was strongly skewed towards brain tissues for all epilepsy phenotypes (Fig. 3d, Supplementary Figs. 10–12).

H3K27ac and H3K4me1 are epigenetic markers associated with regulating gene transcription, suggesting that altered transcription in the dorsolateral prefrontal cortex could be one of the underlying mechanisms of epilepsy. This is further supported by a tissue-specific heritability enrichment analysis (using data from the GTEx Consortium), showing strongest enrichment for genetic generalized epilepsy with genes expressed in Brodmann Area 9 (stratified LD-score regression; $p = 1.56 \times 10^{-6}$), which encompasses the dorsolateral prefrontal cortex (Fig. 3e). These findings further corroborate our TWAS results (using data from the unrelated CommonMind Consortium database), which shows significant associations of epilepsy with gene expression of several genes in the dorsolateral prefrontal cortex (Fig. 2; Supplementary Table 3). Although genetic generalized epilepsy has been regarded as a generalized process, anatomical, electrophysiological, cognitive, and functional imaging studies implicate dysfunction in the frontal lobes[30,31,32,33,34]. Altogether, we have converging evidence from several unrelated methods and databases suggesting epigenetic regulation of gene expression in the dorsolateral prefrontal cortex as a potential pathophysiological mechanism underlying our epilepsy GWAS findings.

Finally, we leveraged the Brainspan database, as implemented in FUMA[35], to assess whether the genes implicated by our GWAS are differentially expressed in the brain at various prenatal and post-natal ages. These analyses were performed for the genes prioritized in any epilepsy phenotype (21 genes), any focal epilepsy subtype (8 genes) or any genetic generalized epilepsy syndrome (15 genes). The results suggest that the expression of

genes associated with focal epilepsy is up-regulated in late-infancy and young adulthood, whereas expression of those genes associated with genetic generalized epilepsy is down-regulated in early childhood and differentially expressed prenatally and at adolescence (Supplementary Fig. 13).

## Enrichment analyses

A previous exome-sequencing study found an association for common epilepsies with ultra-rare variants in known monogenic epilepsy genes[36]. To assess whether common epilepsies are also associated with common variants in monogenic epilepsy genes (see Methods), we pooled the 146 genes that were mapped to our genome-wide significant loci and performed a hypergeometric test. Results illustrated an enrichment of known monogenic epilepsy genes within the genes mapped to our genome-wide significant loci (6 genes overlapped; hypergeometric test: odds ratio [OR] = 8.45, $p = 1.3 \times 10^{-5}$). This enrichment is considerably more significant when limited to the 21 genes with the highest biological priority from Fig 2 (5 genes overlapped; hypergeometric test: OR = 61.4, $p = 9.9 \times 10^{-10}$). We did not find a bias for gene size in our enrichment analyses when using a conservative method to correct for this (see Methods). This suggests that both common and rare variants in monogenic epilepsy genes contribute to common epilepsy susceptibility, corroborating and further extending previous observations[8,37]. Further studies are required to establish whether the signals from common and rare variants are independent of each other.

Using public databases of drug-targets, we found that 13 out of 24 currently licensed anti-epileptic drugs target genes that are implicated in our GWAS. Using the same list of 146 genes as described above, we performed a hypergeometric test which shows a significant enrichment of genes that are known targets of anti-epileptic drugs (8 genes overlapped; hypergeometric test: OR = 19.6, $p = 1.3 \times 10^{-9}$). This enrichment is considerably more significant when limited to the 21 most biologically plausible candidate genes (5 genes overlapped; hypergeometric test: OR = 101.2, $p = 5.7 \times 10^{-11}$). This observation suggests that other drugs that target genes from our GWAS could also have potential for the treatment of epilepsy. The Drug-Gene interaction database

(http://dgidb.org) lists 166 drugs that target biologically prioritized genes from our GWAS (see Supplementary Data 2 for a full list), that may be further investigated for their anti-seizure potential.

Next, we used a complementary approach[38] to search for repurposable drugs. By comparing GWAS-imputed and drug-induced transcriptomes, we predicted drugs capable of rectifying epilepsy-associated gene expression changes (see Methods). Our predictions are enriched with licensed antiepileptic compounds (permutation based $p$-value $<1.0\times10^{-6}$) and with other licensed compounds that have proven antiepileptic efficacy in animal models (permutation based $p$-value$<1.0\times10^{-6}$). We list 30 of our predicted drugs that are licensed for other conditions and have published evidence of efficacy in animal models of epilepsy (Supplementary Table 4).

## Heritability analyses

Twin-based and genetic heritability studies have suggested that while epilepsy is strongly heritable[8,39], there is a substantial missing heritability component[40,41]. We used LDAK to estimate $H^2_{SNP}$; the proportion of heritability that can be attributed to SNPs[42,43,44]. We estimate $H^2_{SNP}$; = 32.1% (95%CI: 29.6–34.5%) for genetic generalized epilepsy and $H^2_{SNP}$; = 9.2% (8.4–10.1%) for focal epilepsy (estimates are on the liability scale, assuming a prevalence of 0.002 and 0.003, respectively) which are consistent with previous estimates[8]. These results indicate that SNPs explain a sizeable proportion of the liability of genetic generalized epilepsy syndromes, but less so for focal epilepsy phenotypes (Fig. 4). To delineate the heritability of the different epilepsy phenotypes, we used LDAK to perform genetic correlation analyses between the different forms. We found evidence for strong genetic correlations between the genetic generalized epilepsies, whereas we found no significant correlations between the focal epilepsies (Fig. 4). Interestingly, we found a significant genetic correlation between JME and lesion-negative focal epilepsy (LDAK genetic correlation: $R^2$=0.46, $p$=$8.77\times10^{-4}$), suggesting either pleiotropy and/or misclassification. It is known that focal EEG features can be seen in JME[45].

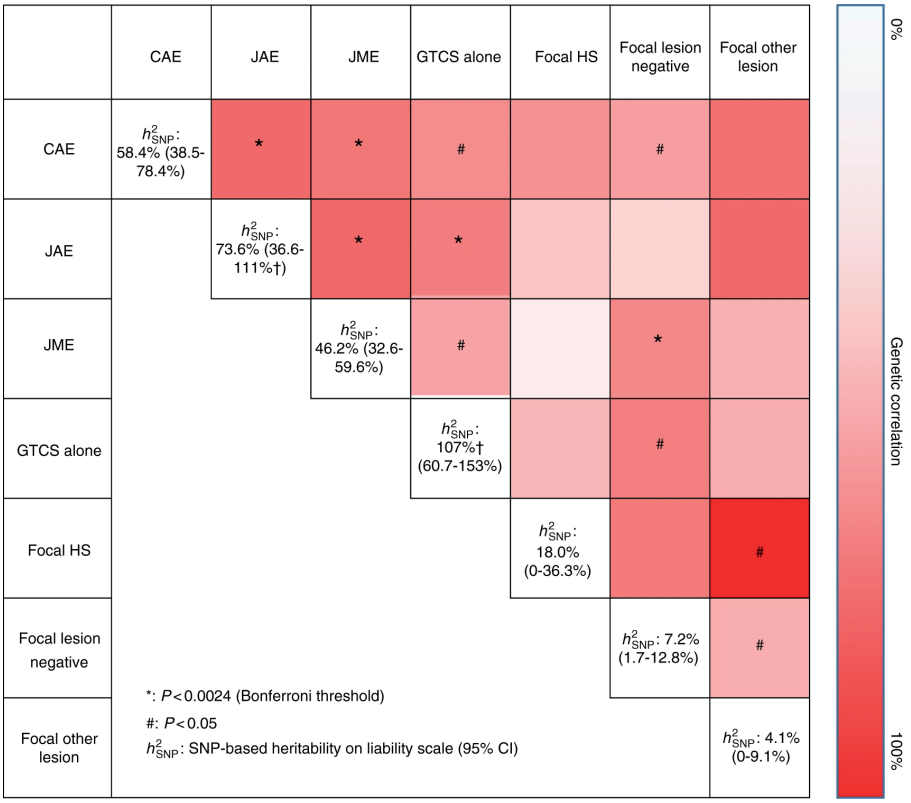



**Figure 4: Heritability estimates and genetic correlations between epilepsy syndromes, calculated using LDAK.** Subjects with a diagnosis of both CAE and JAE were excluded from both phenotypes. The genetic correlation coefficient was calculated with LDAK and is denoted with a color scale ranging from 0% (white) to 100% (red). #P < 0.05; *P < 0.0024 (Bonferroni threshold); $H^2_{SNP}$; SNP-based heritability on liability scale (95% CI); †heritability estimate exceeded 100%, possibly due to small sample size and large SD; CAE - childhood absence epilepsy, JAE - juvenile absence epilepsy, JME - juvenile myoclonic epilepsy, GTCS alone - generalized tonic-clonic seizures alone, focal HS - focal epilepsy with hippocampal sclerosis.

In view of the increasing data on comorbidities with epilepsy, we next used LD-score regression to analyze the genetic correlation between epilepsy and various other brain diseases and traits from previously published GWAS (Fig. 5; see Supplementary Table 5 for values). Perhaps surprisingly, we did not find significant correlations with febrile seizures. Similarly, we did not find any significant genetic correlations between epilepsy and other neurological or psychiatric diseases. However, we did observe significant correlations for

all epilepsy and genetic generalized epilepsy with cognitive ability. We then used the method Multi-Trait Analysis of GWAS (MTAG)[46] to leverage the larger sample size of the genetically correlated GWAS of cognitive ability ($n = 78{,}308$) in order to boost the effective sample size of our all and genetic generalized epilepsy GWAS to 53,244 and 41,515 respectively. Using this approach, we found a novel genome-wide significant locus at 10q24.32 in all epilepsy (MTAG $p = 2.2 \times 10^{-8}$) and genetic generalized epilepsy (MTAG $p = 4.0 \times 10^{-8}$) which encompasses the $K_v$-channel-interacting protein 2 (*KCNIP2*) gene (Supplementary Fig. 14), loss of which is associated with seizure susceptibility in mice[47].

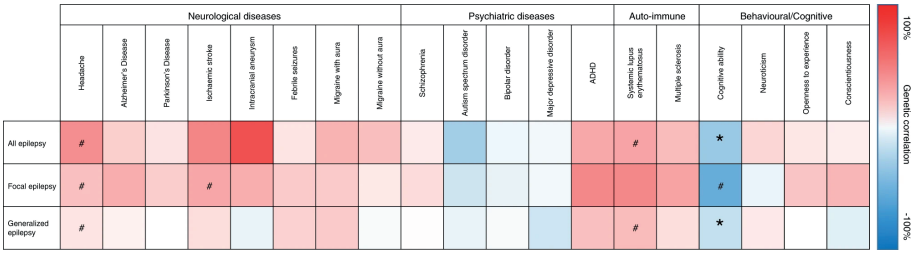



**Figure 5: Genetic correlations of epilepsy with other phenotypes.** The genetic correlation coefficient, calculated using LD-score regression, is denoted with a color scale ranging from -100% (blue) to 100% (red). #: P < 0.05 *P < 0.001 (Bonferroni threshold; 0.05/48)

# Discussion

The increased sample size in this second ILAE Consortium GWAS of common epilepsies has resulted in the detection of 16 risk loci for epilepsy and illustrates how common variants play an important role in the susceptibility of these conditions. But compared to other common neurological diseases our sample size is modest. For example the latest GWAS in schizophrenia considered 36,989 schizophrenia cases and 113,075 controls, resulting in the identification of 108 risk loci[48]. Larger efforts would deliver further insight to the genetic architecture of the common epilepsies.

The majority of the loci are associated with genetic generalized epilepsy. This observation is a welcome partial explanation for the high heritability of genetic generalized epilepsy, in light of the relative lack of rare variant variants

discovered to date. We also show that there is substantial genetic correlation between the generalized syndromes. We speculate that the subtypes share a large part of the genetic susceptibility for generalized epilepsies, with specific modifying factors determining the specific syndrome.

Some syndrome-specific associations were detected, such as the relatively strong signal for *STX1B* in JME, and the association of *GJA1* with focal epilepsy-hippocampal sclerosis. Interestingly, although the association signal for *STX1B* was only significant in the JME analysis, rare pathogenic variants in *STX1B* have been recently found in a spectrum of epilepsies, including genetic epilepsy with febrile seizures plus (GEFS+), genetic generalized epilepsies (including JME), epileptic encephalopathies and even some focal epilepsies[49,50] (Wolking *et al.,* Manuscript submitted (2018). Further, mutations in the gap-junction gene *GJA1* are associated with impaired development of the hippocampus[51] and different expression has been reported in epileptic hippocampal and cortical tissue[52,53]. These findings represent a tantalizing glance of the different biological mechanisms underlying epilepsy syndromes that may guide us to the introduction of genetics for improved diagnosis, prognosis and treatment for these common epilepsies. However, the relatively low sample size of our subtype analysis warrants a conservative interpretation and follow-up with a larger cohort.

At least three association signals are shared between focal epilepsy and genetic generalized epilepsy. The clearest overlapping signal remains the 2q24.3 locus, as we reported previously[15]. However, this association signal is complex and we demonstrate that the locus consists of at least two independent signals (Supplementary Fig. 3). Our Hi-C chromatin analysis suggests the complexity includes levels of functional association to *SCN2A* and *SCN3A*, that are located more distally to the *SCN1A* locus. Mutations in *SCN2A* and more recently *SCN3A* are established monogenic causes of epileptic encephalopathy that, like *SCN1A*, cause dysfunction of the encoded ion-channels, which is believed to disturb the fine balance between neuronal excitation and inhibition. This may involve independent variation that either affects regulation of *SCN1A*, *SCN2A*, or *SCN3A* independently. However, the complex association may also reflect multiple rare risk variations, and large resequencing studies will shed further light on this issue.

The number of association signals we detected and increased power relative to our previous meta analysis[15] allowed us to explore the biological mechanisms behind the observed genetic associations. We show that the signals converge on the dorsolateral prefrontal cortex as the tissue in which most functional effect is observed; this is broadly consistent with the importance of the frontal lobes in generalized epilepsies. Indeed, our analyses of the epigenetic markers H3K27ac and H3K4me1, TWAS, and tissue-specific heritability enrichment analysis all point towards epigenetic regulation of gene expression in the dorsolateral prefrontal cortex as a potential pathophysiological mechanism underlying our epilepsy GWAS findings.

Altogether, we found 16 loci that are associated with the common epilepsies. Our heritability analyses show that collectively, common genetic variants explain a third of the liability for genetic generalized epilepsy. Our analyses suggest that the associated variants are involved in regulation of gene expression in the brain. The 21 biological epilepsy candidate genes implicated by our study have diverse biological functions, and we show that these are enriched for known epilepsy genes and targets of current antiepileptic drugs. Our analyses provide evidence for pleiotropic genetic effects that raise risk for the common epilepsies collectively, as well as variants that raise risk for specific epilepsy syndromes. Determining the shared and unique genetic basis of epilepsy syndromes should be of benefit for further stratification and eventually lead to possible applications for improved diagnosis, prognosis, and treatment. Future studies including pharmacoresponse data, imaging, and other clinical measurements have the potential to further increase the benefit of these studies for people with epilepsy. In combination, these findings further our understanding of the complex genetic architecture of the epilepsies and could provide leads for new treatments and drug repurposing.

# Methods

## Ethics statement

We have complied with all relevant ethical regulations. All study participants provided written, informed consent for use of their data in genetic studies of epilepsy. For minors, written informed consent was obtained from their

parents or legal guardian. Local institutional review boards approved study protocols at each contributing site.



## Cohorts and phenotype definition

A list of the sites included in this study is described in Supplementary Table 6. We classified seizures and epilepsy syndromes according to the classification and terminology outlined by the ILAE[15,54]. For all cases, epilepsy specialists assessed each phenotype at the contributing site. Individuals with epilepsy were initially assigned to one of three phenotypic categories: genetic generalized epilepsy, focal epilepsy, or unclassified epilepsy. Based on EEG, MRI and clinical histories we further classified our cohort into the epilepsy subtypes listed in Supplementary Table 1. We used a combination of population-based datasets as controls. Some control cohorts were screened by questionnaire for neurological disorders. 53.4% of cases were female compared to 51.6% of controls.

## Study design

We conducted a case-control study in subjects of Caucasian, Asian (Han Chinese) and African-American ethnicities. Our primary analyses were structured to test common genetic variants for association with epilepsy according to broad epilepsy phenotypes. We pooled cases from cohorts of the same ethnic group to perform linear mixed model analysis, followed by subsequent meta-analyses of regression coefficients across the three ethnic groups. Our secondary analyses tested for associations with specific syndromes of genetic generalized epilepsy (childhood absence epilepsy, juvenile absence epilepsy, juvenile myoclonic epilepsy, and generalized tonic-clonic seizures alone) and phenotypes of focal epilepsy (lesion negative, focal epilepsy with hippocampal sclerosis, and focal epilepsy with other lesions). The secondary analyses were limited to Caucasian subjects due to sample size. We prioritized the results of the GWAS by incorporating eQTL information, transcriptome-wide analysis, and biological annotation. Finally, we estimated the genetic correlation of epilepsy phenotypes using Linkage-Disequilibrium Adjusted Kinships (LDAK).

## Genotyping

The EpiPGX samples were genotyped at deCODE Genetics on Illumina OmniExpress-12 v1.1 and OmniExpress-24 v1.1 single nucleotide polymorphism (SNP) arrays. The EPGP samples were genotyped on Illumina HumanCore beadchips at Duke University, North Carolina. The remainder of the samples were genotyped on various SNP arrays, as previously published[15].

## Genotyping quality control and imputation

Quality control of genotyping was performed separately for each cohort using PLINK 1.9[55]. Each genotype cohort was temporarily merged with a genetically similar reference population from the 1000 Genomes Project (CEU, CHB, or YRI). A test for Hardy–Weinberg equilibrium (HWE) was performed and SNPs significant at $p < 1 \times 10^{-10}$ were removed. All samples and all SNPs with missing genotype rate >0.05 and all SNPs with minor allele frequency (MAF) <0.01 were removed. Next, we pruned SNPs using the PLINK --indep-pairwise command (settings: window size 100 kb, step size 25, $R^2 > 0.1$). Using this subset of SNPs, we removed samples with outlying heterozygosity values (>5 SD from the median of the whole cohort). Identity by descent (IBD) was calculated in PLINK to remove sample duplicates (>0.9 IBD) and to identify cryptic relatedness. We removed one from each sample pair with IBD>0.1875, with the exception of the EPGP familial epilepsy cohort. Subjects were removed if sex determined from X-chromosome genotype did not match reported gender. Array-specific maps were used to update all SNPs positions and chromosome numbers to the Genome Reference Consortium Human Build 37 (GRCh37), and remove all A/T and C/G SNPs to avoid strand issues. We applied pre-imputation checks according to scripts available on the website of Will Rayner of the Wellcome Trust Centre for Human Genetics (www.well.ox.ac.uk/~wrayner/tools/) to remove SNPs with allele frequencies deviating >20% from the frequency in the Haplotype Reference Consortium. Samples were submitted to the Sanger Imputation Service (https://imputation.sanger.ac.uk/)[56]. We selected the Human Reference Consortium (release 1.1; n = 32470) dataset as reference panel for Caucasian and Asian datasets and the African Genome

Resources ($n = 4956$) for the African-American datasets. Post-imputation quality control filters were applied to remove SNPs within each imputed cohort with an imputation info score <0.9 or HWE $p$<1e-6. Imputed genotype dosages with a minimum certainty of 0.9 per subject were converted to hard-coded PLINK format after which all samples were pooled into a single cohort. We performed a principal components analysis using GCTA. From the PCA results we stratified our subjects into three broad ethnic groups (Caucasian, Asian and African) while removing extreme outliers. After stratifying by ethnicity, we removed SNPs with HWE $p$ < 1e-6, call rate <0.95 or MAF<0.01. In total 816 subjects out of 45705 subjects were filtered out by quality control procedures, leaving 44889 subjects for analyses.

## Study power

We estimated using PGA[57] that the study had 80% power to detect a genetic predictor of relative risk for all epilepsy (approximated to odds ratio) ≥1.45 with MAF = 1% and an alpha level of $5 \times 10^{-8}$. We estimated that our meta-analyses had 80% power to detect genome-wide significant SNPs of MAF = 1% with relative risks ≥1.5 and ≥1.8, for focal and generalized epilepsy respectively (see Supplementary Figure 15). Our analysis of generalized epilepsy sub-phenotypes had 80% power to detect genome-wide significant SNPs of MAF = 1% with relative risks ≥2.6, ≥3.3, and ≥2.4 for CAE, JAE, and JME respectively. Our analysis of focal epilepsy sub-phenotypes had 80% power to detect genome-wide significant SNPs of MAF = 1% with relative risks ≥1.9, ≥2.8, and ≥1.9 for focal epilepsy lesion negative, focal epilepsy with hippocampal sclerosis and focal epilepsy with lesion other than hippocampal sclerosis, respectively.

## Statistical analyses

Association analyses were conducted within the three ethnic subgroups using a linear mixed model in BOLT-LMM[58]. A subset of SNPs, used to correct for (cryptic) relatedness and population stratification by BOLT-LMM, were derived by applying SNP imputation info score >0.99, MAF >0.01, call rate >0.99 before pruning the remaining variants using LDAK with a window size of 1Mb and $R^2 > 0.2$[43]. All analyses included gender as a

covariate and the threshold for statistical significance was set at $5 \times 10^{-8}$. We compared $\chi^2$ values of the BOLT-LMM output between all pairs of SNPs in high LD ($R^2 > 0.4$) and removed pairs of SNPs with extreme $\chi^2$ differences using a formula that scales exponentially with magnitude of $\chi^2$ and LD: $\chi^2$ difference cutoff = $\frac{3*\sqrt{\frac{SNP1-\chi2+SNP2-\chi2}{2}}}{(R^2)^2}$ ; where *SNP1*- $\chi^2$ and *SNP2*- $\chi^2$ are the $\chi^2$-statistic of the two SNPs in each pair and $R^2$ is their squared correlation (LD). We tested the homogeneity of all SNPs by splitting the pooled cohort into 13 distinct clusters of ethnically matched cases and controls and removed SNPs exhibiting significant heterogeneity of effect ($P_{het} < 1 \times 10^{-8}$). Fixed effects, trans-ethnic meta-analyses were conducted using the software package METAL[59]. Manhattan plots for all analyses were created using qqman. Considering that our study had unequal case-control ratios, we calculated the effective sample size per ethnicity using the formula recommended by METAL: $N_{eff} = 4/(1/N_{cases} + 1/N_{ctrls})$. Since >95% of all cases were Caucasian, we included all SNPs that were present in at least the Caucasian dataset (~5 million).

Conditional association analysis was performed with PLINK on loci containing significant SNPs to establish whether other genetic variants in the region (500 Kb upstream and downstream) were independently associated with the same phenotype. The conditional threshold for significance was set at $2 \times 10^{-5}$, based on approximately 2500 imputed variants per 1MB region.

## Assessment of inflation of the test statistic

Potential inflation of the test statistic was assessed per ethnicity and phenotype by calculating the genomic inflation factor ($\lambda$; the ratio of the median of the empirically observed distribution of the test statistic to the expected median) and the mean $\chi^2$. Since $\lambda$ is known to scale with sample size, we also calculated the $\lambda_{1000}$, i.e $\lambda$ corrected for an equivalent sample size of 1000 cases and 1000 controls[60]. We observed some inflation of the test statistic ($\lambda > 1$) across the different phenotypes (Supplementary Table 7), suggesting either polygenicity or confounding due to population stratification or cryptic relatedness. Therefore, we applied LD score regression[61], estimating LD scores using matched populations from the 1000 GP (EUR for Caucasians

($n = 669$), AFR for African-Americans and EAS for Asians). These LDSC results suggested that inflation of the test statistic was primarily due to polygenicity for most analyses (Supplementary Table 7). Only the Caucasian focal and all epilepsy analyses had LDSC intercepts >1.1, suggesting confounding or an incomplete match of the LD-score reference panel. Our focal and all epilepsy analyses included cohorts from various Caucasian ethnicities, including Finnish and Italian focal epilepsy cohorts, and it has been shown that LD differs considerably between Finnish and Italian populations[61]. Therefore, we consider an incomplete match of the LD-score reference panel the most likely cause of the observed inflation, since we used a mixed-model analysis that corrects for population stratification and cryptic relatedness[58].

## Gene mapping and biological prioritization

Genome-wide significant loci of all phenotypes were mapped to genes in and around these loci using FUMA[35]. Genome-wide significant loci were defined as the region encompassing all SNPs with $P < 10^{-4}$ that were in LD ($R^2 > 0.2$) with the lead SNP (i.e. the SNP with the lowest $P$-value in the locus with $P < 5 \times 10^{-8}$). Positional mapping was performed to map genes that were located within 250 kb of these loci. Additionally, we mapped genes that were farther than 250 kb away from the locus using chromatin interaction data to identify genes that show a significant 3D interaction ($P_{FDR} < 10^{-6}$) between their promoter and the locus, based on Hi-C data from dorsolateral prefrontal cortex, hippocampus, and neural progenitor cells[62]. This resulted in a total of 146 mapped genes across all phenotypes, of which some genes (e.g. *SCN1A*) were associated with multiple epilepsy phenotypes.

We next devised various prioritization criteria to prioritize the most likely biological candidate genes out of the 146 mapped genes, similar to previous studies[17,18,63], based on the following 6 criteria:

1. A significant correlation between the epilepsy phenotype and expression of the gene, as assessed with a transcriptome-wide association study (TWAS). Default settings of the FUSION software package[64] were used to impute gene-expression based on our GWAS summary statistics and RNA-sequencing data from dorsolateral prefrontal cortex tissue ($n = 452$, CommonMind Consortium[65]), after which the association

between the epilepsy phenotype with gene-expression was calculated. It was possible to test the TWAS expression association for 53 out of our 146 mapped genes, since only the expression of these 53 genes was significantly heritable (heritability $p$-value <0.01, as suggested by Gusev et al.[64]). We set a Bonferroni corrected $p$-value threshold of 0.05/53 = 0.00094 to define significant TWAS associations.

2. Genes for which a SNP in the genome-wide significant locus (as defined above) is a significant cis-eQTL (Bonferroni corrected $P < 8 \times 10^{-10}$)[23] based on data from the ROS and MAP studies, which includes RNA-sequencing data from 494 dorsolateral prefrontal cortex tissues[23].
3. The gene is preferentially expressed in the brain. This was assessed by using gene-expression data from all 53 tissues of the Gene-Tissue expression (GTEx) Consortium[66]. Genes were considered to be preferentially expressed in the brain when the average expression in all brain tissues was higher than the average expression in non-brain tissues.
4. Genes for which a SNP in the genome-wide significant locus (as defined above) is a missense variant, as annotated by ENSEMBL[67].
5. Genes prioritized by protein-protein interaction network, as calculated using the default settings of DAPPLE[68], which utilizes protein–protein interaction data from the InWeb database[69]. The 146 genes implicated by our GWAS were input after which DAPPLE assessed direct and indirect physical interactions to create a protein-interaction network. Next, DAPPLE assigned a significance score to each gene according to several connectivity parameters; genes with a corrected $P < 0.05$ were considered to be prioritized by DAPPLE.
6. Genes for which a nervous system or behavior/neurological phenotype was observed in knockout mice. Phenotype data of knockout mice was downloaded from the Mouse Genome Informatics database (http://www.informatics.jax.org/) on 17 January 2018 and nervous system (phenotype ID: MP:0003631) and behavior/neurological phenotype (MP:0005386) data were extracted.

All 146 genes were scored based on the number of criteria met (range 0–6 with an equal weight of 1 per criterion), see Supplementary Data **1** for a full list. We considered the gene(s) with the highest score in each locus as the most likely biological epilepsy candidate gene. Multiple genes in a locus were

selected if they had an equally high score whilst no genes were selected in a locus if all genes within it had a score <2, similar to previous studies[17,18].



## Gene co-expression analysis for epilepsy with brain-coX

In silico gene prioritization was performed using brain-coX[25]. brain-coX uses a compendium of seven large-scale normal brain gene expression data resources to identify co-expressed genes with a set of given genes (known, or putative, disease causing genes) likely to encapsulate gene expression networks involved in disease. This approach can identify, and thus leverage networks that are not currently known and not present in available resources such as PPI networks and is a complementary approach to these. We used a set 102 monogenic epilepsy genes (Supplementary Table 8) as the set of known disease genes. An FDR of 0.2 was used to identify genes that significantly co-express with the known set of genes. Prioritization in at least three datasets at an FDR of 0.2 led to a specificity of 0.9[25].

In the first analysis we used a set of 16 candidate epilepsy genes identified by the GWAS analysis and prioritized using additional methods (Fig. 2). These excluded any genes already included in the set of known epilepsy genes (Supplementary Table 8). Supplementary Fig. 9 shows the gene co-expression pattern using the weighted average gene co-expression across all seven datasets for candidate genes from the GWAS that show significant gene co-expression with any of the 102 known epilepsy genes.

In the second analysis we used the set of all the 146 candidate genes identified in the GWAS analysis (Supplementary Data 1). Only 140 of these were identified as having available expression data in the gene expression resources. Many genes showed some evidence of gene co-expression but few showed co-expression in more than 2 datasets (18 out of 140). BCL11A (6) and GJA (6) remain the most robust candidate genes co-expressed with known epilepsy genes. The complete results are shown in Supplementary Table 9.

## Functional annotations

We annotated all genome-wide significant SNPs ($p < 5 \times 10^{-8}$) from all phenotypes using the Variant Effect Predictor of ENSEMBL[67] and the

RegulomeDB database[28]. We annotated chromatin states using epigenetic data from the NIH Roadmap Epigenomics Mapping Consortium[70] and ENCODE[71]. We used FUMA[35] to annotate the minimum chromatin state (i.e. the most active state) across 127 tissues and cell types for each SNP, similar to a previous study[27].

## Heritability enrichment of epigenetic markers and gene-expression

We used stratified LD-score regression[72] to assess tissue-specific heritability enrichment of epigenetic markers in 88 tissues, using standard procedures[29]. We used the same settings and pre-calculated weights that accompanied the paper by Finucane et al. to calculate the heritability enrichment of all epilepsy, focal epilepsy and generalized epilepsy, based on epigenetic data of 6 chromatin markers in 88 tissues from the Roadmap Consortium and gene-expression data in 53 tissues from the GTEx Consortium.

## Enrichment analyses

Hypergeometric tests were performed with R (version 3.4.0) to assess whether the genes mapped to genome-wide significant loci and the subset of prioritized biological epilepsy genes (see above) were enriched for: (i) known monogenic epilepsy genes ($n = 102$) and (ii) known anti-epileptic drug target genes ($n = 64$), relative to the rest of the protein-coding genes in the genome ($n = 19180$). We supplemented the list of 43 known dominant epilepsy genes[36] with an additional 59 monogenic epilepsy genes from the GeneDX comprehensive epilepsy panel (www.genedx.com). We compiled the list of drug target genes from[73], supplemented with additional FDA & EMA licensed AEDs. The full list of gene targets considered in each analysis are listed in Supplementary Tables 8 and 10.

## Enrichment analyses corrected for gene size

Brain expressed genes are known to be larger in size than non-brain expressed genes. To assess whether gene size could be a cause of bias for our enrichment analyses, we first assessed whether the size of the

genes mapped in our analyses was different than non-mapped genes in the genome. We found that the size of the 146 genes mapped to genome-wide significant loci was 65.6 kb, whereas the average gene size of all other protein-coding genes is on average 62.2 kb, suggesting there is no strong bias towards preferentially mapping loci to small or large genes.

We also observed that the 102 established monogenic epilepsy genes are on average 2.44 times longer than non-epilepsy genes (152.0 kb vs 62.2 kb). As a conservative approach to correct for this size difference, we have used the Wallenius' noncentral hypergeometric distribution, as implemented in the R-package 'BiasedUrn'. Using this distribution, we repeated our hypergeometric analyses under the conservative assumption of a 2.42 times increased likelihood of mapping epilepsy genes as opposed to non-epilepsy genes. Using this distribution, the 146 genes that were mapped to genome-wide significant loci were significantly enriched for monogenic epilepsy genes (Wallenius' noncentral hypergeometric test $p = 8.3\times10^{-3}$). When limiting our results to the 21 biological prioritized genes, the enrichment of monogenic epilepsy genes became more significant (Wallenius' noncentral hypergeometric distribution $p = 5.3\times10^{-4}$).

Similarly, we observed that the targets of AEDs are on average 2.43 times longer than non-AED target genes (151.8 kb vs 62.4 kb). When correcting for this gene-size difference under the assumption of a 2.43 times increased likelihood of mapping our genome-wide significant loci to AED target genes, we find that the 146 mapped genes were significantly enriched for AED target genes (Wallenius' noncentral hypergeometric test $p = 1.7\times10^{-5}$). When limiting our results to the 21 biological prioritized genes, the enrichment of AED target genes became more significant (Wallenius' noncentral hypergeometric test $p = 1.0\times10^{-8}$).

## Connectivity mapping

Connectivity mapping was performed using our GWAS results in order to identify drugs which can potentially be repurposed for the treatment of epilepsy, enabling significant savings in the time and cost of antiepileptic drug development. Recently, So et al. identified candidate drugs that could

be repurposed for the treatment of schizophrenia by using GWAS results to impute the gene-expression changes associated with the disease and, then, identifying drugs that change gene-expression in the opposite direction in cell lines[38]. Interestingly, the set of candidate drugs they identified was significantly enriched with antipsychotics. We adopted a similar strategy.

Gene-expression changes associated with epilepsy were imputed from the all epilepsy GWAS summary statistics using the FUSION software package[64] and dorsolateral prefrontal cortex tissue RNA-sequencing data ($n = 452$, CommonMind Consortium[65]). We calculated z-scores for the association between epilepsy and changes in expression of all 5261 significantly heritable genes, using default settings of the FUSION software package as described above[64]. The top 10% of the gene-expression changes most strongly associated with epilepsy were used to construct the disease signature. Then, we identified drugs that change gene-expression in the opposite direction in cell lines, using the Combination Connectivity Mapping bioconductor package and the Library of Integrated Network-Based Cellular Signatures (LINCS) data[74]. This package utilizes cosine distance as the (dis)similarity metric[75,76]. A higher (more negative) cosine distance value indicates that the drug induces gene-expression changes more strongly opposed to those associated with the disease. A lower (more positive) cosine distance value indicates that the drug induces gene-expression changes more similar to those associated with the disease. In the LINCS dataset, some drugs have been profiled in more than one cell line, concentration, and time-point. For such drugs, the highest absolute cosine distance, whether positive or negative, was selected, as this value is less likely to occur by chance. The output of this analysis comprised 24,051 drugs or 'perturbagens', each with a unique cosine distance value.

To demarcate the set of drugs predicted to significantly reverse epilepsy-associated gene-expression changes, the threshold of statistical significance for cosine distance values was determined. For this, we performed 100 permutations of the disease gene-expression $z$-scores and compared them to drug gene-expression signatures. We combined the distribution of cosine distance values across all permutations, such that the null distribution was derived from 2,405,100 cosine distance values under $H_0$.

The cosine distance value corresponding to $\alpha$ of 0.05 was −0.386. Of the drugs with a cosine distance less than −0.386, thirty were experimentally-validated drug repurposing candidates from the Prescribable Drugs with Efficacy in Experimental Epilepsies (PDE3) database—a recently published systematic and comprehensive compilation of licenced drugs with evidence of antiepileptic efficacy in animal models[77]. We determined whether this is more than expected by chance, by creating 1,000,000 random drug-sets of the same size as drugs with a significant cosine distance. Next, we counted the number of subsets containing an equal or higher number of experimentally-validated drug repurposing candidates, as those found within drugs with a significant cosine distance. This permutation-based $p$-value was $1.0 \times 10^{-6}$.

Supplementary Table 4 lists the 30 candidate re-purposable drugs that are predicted to significantly reverse epilepsy-associated gene-expression changes, have published evidence of antiepileptic efficacy in animal models, and are already licensed for the treatment of other human diseases. Of this list, 22 drugs have corroborated evidence of antiepileptic efficacy from multiple published studies or multiple animal models. For each drug, we list the studies demonstrating antiepileptic efficacy in animal models, the animal models used, and the licensed indication(s).

## Validation of connectivity mapping results

Validation of the connectivity mapping results was performed using two non-overlapping sets of drugs with known antiepileptic efficacy: (1) a set of 'clinically-effective' drugs that have antiepileptic efficacy in people, and (2) a set of 'experimentally-validated' drugs that have antiepileptic efficacy in animal models. For the clinically-effective drug-set, we used the names of all recognized antiepileptic drugs, as listed in category N03A of the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) Classification System, and of benzodiazepines and their derivatives (ATC codes N05BA and N05CD), and of barbiturates (ATC code N05CA), as these drugs are known to have antiepileptic efficacy in people. For the experimentally-validated drug-set, we extracted drug names from the PDE3 database[77].

We determined whether, in our results, clinically effective drugs are ranked higher than expected by chance. The median rank of all drugs was 12,026. The median rank of clinically effective drugs was 3725. Hence, the median rank of clinically-effective drugs was 8301 positions higher than that of all drugs. A permutation-based $p$-value was determined by calculating the median ranks of 1,000,000 random drug-sets, each equal in size to the number of clinically effective drugs in the LINCS database. This permutation-based $p$-value was $<1.0 \times 10^{-6}$. Similarly, we determined whether, in our results, experimentally-validated drugs are ranked higher than expected by chance. The median rank of experimentally-validated drugs was 6372. Hence, the median rank of experimentally-validated drugs was 5654 positions higher than that of all drugs. A permutation-based $p$-value was determined by calculating the median ranks of 1,000,000 random drug-sets, each equal in size to the number of experimentally-validated drug repurposing candidates in the LINCS database. This permutation-based $p$-value was $<1.0 \times 10^{-6}$.

## Heritability analysis

Linkage-Disequilibrium Adjusted Kinships (LDAK[42,43]) was used to calculate SNP-based heritability of all epilepsy phenotypes. Since these analyses require homogeneous cohorts, only Caucasian subjects (which represent >95% of epilepsy cases) were used for these analyses. SNP based heritabilities (h2oho2) were converted to liabilitv scale heritability estimates (h2LhL2) using the formula: $h_L^2 = h_o^2 * \frac{K^2(1-K)^2}{p(1-p)} * Z^2$, where $K$ is the disease prevalence, $p$ is the proportion of cases in the sample, and $Z$ is the standard normal density at the liability threshold. We estimated disease prevalence based on a combination of previous studies[8,78,79] (Supplementary Table 11). Although prevalence estimates vary between studies, the h2LhL2 estimate has been shown to be fairly robust to such differences[8]. Similarly, we have modeled h2LhL2 using half and double of our prevalence estimates which lead to h2LhL2 estimates that varied between 0.4 and 11% (Supplementary Table 11). In addition, we compared the heritability estimates from LDAK with the alternative methods BOLT-REML[80] and LDSC[58] (Supplementary Table 12). Next, LDAK was used to calculate the genetic correlation between the 7 epilepsy subtypes. Subjects with a diagnosis of both CAE and JAE were excluded from heritability and genetic correlation analyses.

We computed the genetic correlation between all, focal and genetic generalized epilepsy with other brain diseases and traits using LDSC, as implemented in LD hub[81]. LD hub is a centralized database that contains publicly available GWAS summary statistics from various diseases and traits. We selected published GWAS of psychiatric, neurological, auto-immune diseases with known brain involvement and cognitive/behavioral traits from LD hub. We contacted the authors of published GWAS to provide us with summary statistics when no summary statistics were available on LDhub or when a more recent GWAS of a disease/trait was published that was not included on LDhub. The Caucasian subset of our data was used for all analyses and only other GWAS with primarily Caucasian subjects were included in our analyses. We used a conservative Bonferroni correction to assess significance of genetic correlations ($p = 0.05/48 = 0.001$).

Multi-trait analysis of GWAS (MTAG)[46] was used with default settings to increase the effective sample size from our Caucasian all and generalized epilepsy GWAS by pairing it with the significantly correlated GWAS on cognitive ability (as assessed above) with a larger sample size ($n$=78,307). MTAG utilizes the fact that estimations of effect size and standard error of a primary GWAS, in this case epilepsy, can be improved by matching them to a genetically correlated secondary GWAS, in this case cognitive ability.

## Supporting information

Additional supporting information may be found at: https://tinyurl.com/4j3srm5j

## Acknowledgements

We are grateful to the patients and volunteers who participated in this research. We thank the following clinicians and research scientists for their contribution through sample collection (cases and controls), data analysis, and project support: Geka Ackerhans, Muna Alwaidh, R E Appleton, Willem Frans Arts, Guiliano Avanzini, Paul Boon, Sarah Borror, Kees Braun, Oebele Brouwer, Hans Carpay, Karen Carter, Peter Cleland, Oliver C Cockerell, Paul Cooper, Celia Cramp, Emily de los Reyes, Chris French, Catharine Freyer,

William Gallentine, Michel Georges, Peter Goulding, Micheline Gravel, Rhian Gwilliam, Lori Hamiwka, Steven J Howell, Adrian Hughes, Aatif Husain, Monica Islam, Floor Jansen, Mary Karn, Mark Kellett, Ditte B Kjelgaard, Karl Martin Klein, Donna Kring, Annie WC Kung, Mark Lawden, Jo Ellen Lee, Benjamin Legros, Leanne Lehwald, Edouard Louis, Colin HT Lui, Zelko Matkovic, Jennifer McKinney, Brendan McLean, Mohamad Mikati, Bethanie Morgan-Followell, Wim Van Paesschen, Anup Patel, Manuela Pendziwiat, Marcus Reuber, Richard Roberts, Guy Rouleau, Cathy Schumer, B Sharack, Kevin Shianna, NC Sin, Saurabh Sinha, Laurel Slaughter, Sally Steward, Deborah Terry, Chang-Yong Tsao, TH Tsoi, Patrick Tugendhaft, Jaime-Dawn Twanow, Jorge Vidaurre, Sarah Weckhuysen, Pedro Weisleder, Kathleen White, Virginia Wong, Raju Yerra, Jacqueline Yinger and all contributing clinicians from the Department of Clinical and Experimental Epilepsy at the National Hospital for Neurology and Neurosurgery and UCL Institute of Neurology. Data generated as part of the EPIGEN Consortium was included in this study. We would like to thank Dr. Weihua Meng (University of Dundee), Dr. Mark Adams (University of Edinburgh) and Dr. Ynte Ruigrok (UMC Utrecht), Dr. Bjarke Feenstra (Statens Serum Institut, Denmark), Dr. Risto Kayanne (Institute for Molecular Medicine Finland) and the International Headache Genetics Consortium for providing GWAS summary statistics for their respective cohorts. Ischemic stroke summary statistics were accessed through the ISGC Cerebrovascular Disease Knowledge Portal. We would like to thank Dr. Weihua Meng (University of Dundee), Dr. Mark Adams (University of Edinburgh) and Dr. Ynte Ruigrok (UMC Utrecht), Dr. Bjarke Feenstra (Statens Serum Institut, Denmark), Dr. Risto Kayanne (Institute for Molecular Medicine Finland), and the International Headache Genetics Consortium for providing GWAS summary statistics for their respective cohorts. We would like to thank the Ming Fund for providing funding for Re.St. M.McC. has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 751761. This work was in part supported by an award by a Translational Research Scholars award from the Health Research Board of Ireland (C.D.W.), by research grants from Science Foundation Ireland (SFI) (16/RC/3948 and X) and co-funded under the European Regional Development Fund and by FutureNeuro industry partners. Further

funding sources include: Wellcome Trust (grant 084730); Epilepsy Society, UK, NIHR (08-08-SCC); GIHE: NIH R01-NS-49306-01 (R.J.B.); NIH R01-NS-053998 (D.H.L); GSCFE: NIH R01-NS-064154-01 (R.J.B. and Ha.Ha.); NIH: UL1TR001070, Development Fund from The Children's Hospital of Philadelphia (Ha.Ha.); NHMRC Program Grant ID: 1091593 (S.F.B., I.E.S., K.L.O., and K.E.B.); The Royal Melbourne Hospital Foundation Lottery Grant (S.P.); The RMH Neuroscience Foundation (T.J.O'B.); European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 279062 (EpiPGX) and 602102, Department of Health's NIHR Biomedical Research Centers funding scheme, European Community (EC: FP6 project EPICURE: LSHM-CT-2006-037315); German Research Foundation (DFG: SA434/4-1/4-26-1 (Th.Sa.), WE4896/3-1); EuroEPINOMICS Consortium (European Science Foundation/DFG: SA434/5-1, NU50/8-1, LE1030/11-1, HE5415/3-1 (Th.Sa., P.N., H.L., I.H.), RO 3396/2-1); the German Federal Ministry of Education and Research, National Genome Research Network (NGFNplus/EMINet: 01GS08120, and 01GS08123 (Th.Sa., H.L.); IntenC, TUR 09/I10 (Th.Sa.)); The Netherlands National Epilepsy Fund (grant 04-08); EC (FP7 project EpiPGX 279062). Research Grants Council of the Hong Kong Special Administrative Region, China project numbers HKU7623/08M (S.S.C, P.K., L.W.B., P.C.S), HKU7747/07M (S.S.C., P.C.S.) and CUHK4466/06M (P.K., L.B). Collection of Belgian cases was supported by the Fonds National de la Recherche Scientifique, Fondation Erasme, Université Libre de Bruxelles. GlaxoSmithKline funded the recruitment and data collection for the GenEpA Consortium samples. We acknowledge the support of Nationwide Children's hospital in Columbus, Ohio, USA. The Wellcome Trust (WT066056) and The NIHR Biomedical Research Centres Scheme (P31753) supported UK contributions. Further support was received through the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Contract: N01HD33348). The project was also supported by the popgen 2.0 network through a grant from the German Ministry for Education and Research (01EY1103). Parts of the analysis of this work were performed on resources of the High Performance Center of the University of Luxembourg and Elixir-Luxembourg. The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German

Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. The International League Against Epilepsy (ILAE) facilitated the Consortium through the Commission on Genetics and by financial support; however, the opinions expressed in the manuscript do not necessarily represent the policy or position of the ILAE.

## Authors

Bassel Abou-Khalil, Pauls Auce, Andreja Avbersek, Melanie Bahlo, David J. Balding, Thomas Bast, Larry Baum, Albert J. Becker, Felicitas Becker, Bianca Berghuis, Samuel F. Berkovic, Katja E. Boysen, Jonathan P. Bradfield, Lawrence C. Brody, Russell J. Buono, Ellen Campbell, Gregory D. Cascino, Claudia B. Catarino, Gianpiero L. Cavalleri, Stacey S. Cherny, Krishna Chinthapalli, Alison J. Coffey, Alastair Compston, Antonietta Coppola, Patrick Cossette, John J. Craig, Gerrit-Jan de Haan, Peter De Jonghe, Carolien G. F. de Kovel, Norman Delanty, Chantal Depondt, Orrin Devinsky, Dennis J. Dlugos, Colin P. Doherty, Christian E. Elger, Johan G. Eriksson, Thomas N. Ferraro, Martha Feucht, Ben Francis, Andre Franke, Jacqueline A. French, Saskia Freytag, Verena Gaus, Eric B. Geller, Christian Gieger, Tracy Glauser, Simon Glynn, David B. Goldstein, Hongsheng Gui, Youling Guo, Kevin F. Haas, Hakon Hakonarson, Kerstin Hallmann, Sheryl Haut, Erin L. Heinzen, Ingo Helbig, Christian Hengsbach, Helle Hjalgrim, Michele Iacomino, Andrés Ingason, Jennifer Jamnadas-Khoda, Michael R. Johnson, Reetta Kälviäinen, Anne-Mari Kantanen, Dalia Kasperavičiūte, Dorothee Kasteleijn-Nolst Trenite, Heidi E. Kirsch, Robert C. Knowlton, Bobby P. C. Koeleman, Roland Krause, Martin Krenn, Wolfram S. Kunz, Ruben Kuzniecky, Patrick Kwan, Dennis Lal, Yu-Lung Lau, Anna-Elina Lehesjoki, Holger Lerche, Costin Leu, Wolfgang Lieb, Dick Lindhout, Warren D. Lo, Iscia Lopes-Cendes, Daniel H. Lowenstein, Alberto Malovini, Anthony G. Marson, Thomas Mayer, Mark McCormack, James L. Mills, Nasir Mirza, Martina Moerzinger, Rikke S. Møller, Anne M. Molloy, Hiltrud Muhle, Mark Newton, Ping-Wing Ng, Markus M. Nöthen, Peter Nürnberg, Terence J. O'Brien, Karen L. Oliver, Aarno Palotie, Faith Pangilinan, Sarah Peter, Slavé Petrovski, Annapurna

Poduri, Michael Privitera, Rodney Radtke, Sarah Rau, Philipp S. Reif, Eva M. Reinthaler, Felix Rosenow, Josemir W. Sander, Thomas Sander, Theresa Scattergood, Steven C. Schachter, Christoph J. Schankin, Ingrid E. Scheffer, Bettina Schmitz, Susanne Schoch, Pak C. Sham, Jerry J. Shih, Graeme J. Sills, Sanjay M. Sisodiya, Lisa Slattery, Alexander Smith, David F. Smith, Michael C. Smith, Philip E. Smith, Anja C. M. Sonsma, Doug Speed, Michael R. Sperling, Bernhard J. Steinhoff, Ulrich Stephani, Remi Stevelink, Konstantin Strauch, Pasquale Striano, Hans Stroink, Rainer Surges, K. Meng Tan, Liu Lin Thio, G. Neil Thomas, Marian Todaro, Rossana Tozzi, Maria S. Vari, Eileen P. G. Vining, Frank Visscher, Sarah von Spiczak, Nicole M. Walley, Yvonne G. Weber, Zhi Wei, Judith Weisenberg, Christopher D. Whelan, Peter Widdess-Walsh, Markus Wolff, Stefan Wolking, Wanling Yang, Federico Zara & Fritz Zimprich

## Contributions

Data analysis: protocol development and main analyses: G.L.C., B.P.C.K., Ro.Kr. (data management), De.La., C.L., M.McC. (co-lead analyst), N.M., D.S., and Re.St. (co-lead analyst); analysis coordination: G.L.C., B.P.C.K., and D.S.; data preparation, imputation and quality control: M.B., D.J.B., L.B., J.P.B., R.J.B., G.L.C., S.S.C., A.J.C., C.G.F.deK., S.F., D.B.G., H.G., Y.G., Ha.Ha., E.L.H., I.H., A.I., D.K., B.P.C.K., Ro.Kr., De.La., C.L., I.L-C., A.M., M.McC., N.M., P.-W.N., P.N., Sa.Pe., Sl.Pe., Th.Sa., P.C.S., A.S., D.S., Re.St., Z.W., C.D.W., and Fe.Za.; analysis review: D.J.B., and Ha.Ha. Writing committee: S.F.B., G.L.C., B.P.C.K., M.McC. (co-wrote first draft), and Re.St. (co-wrote first draft). Strategy committee: L.B., S.F.B., R.J.B., G.L.C., Ha.Ha., E.L.H., M.R.J., Re.Kä., B.P.C.K., Ro.Kr., P.K., H.L., I.L-C., T.J.O'B., and S.M.S. Phenotyping committee: C.D., D.J.D., W.S.K., P.K., D.H.L., A.G.M., M.R.S., and P.S. Governance committee: S.F.B., Al.Co., A.-E.L., and D.H.L. Patient recruitment and phenotyping: B.A.-K., P.A., A.A., T.B., A.J.B., F.B., B.B., S.F.B., R.J.B., E.C., G.D.C., C.B.C., K.C., An.Co., P.C., J.J.C., G-J.deH., P.De.J., N.D., C.D., O.D., D.J.D., C.P.D., C.E.E., T.N.F., M.F., B.F., J.A.F., V.G., E.B.G., T.G., S.G., K.F.H., K.H., S.H., I.H., C.H., He.Hj., M.I., J.J-K., M.R.J., Re.Kä., A.-M.K., D.K.-N.T., H.E.K., R.C.K., M.K., W.S.K., Ru.Ku., P.K., H.L., Di.Li., W.D.L., I.L-C., D.H.L., A.G.M., T.M., M.M., R.S.M.,

H.M., M.N., P.-W.N., T.J.O'B., An.Po., M.P., R.R., S.R., P.S.R., E.M.R., F.R., J.W.S., Th.Sa., Th.Sc., S.C.S., C.J.S., I.E.S., B.S., S.S., J.J.S., G.J.S., S.M.S., L.S., D.F.S., M.C.S., P.E.S., A.C.M.S., M.R.S., B.J.S., U.S., P.S., H.S., Ra.Su., K.M.T., L.L.T., M.T., R.T., M.S.V., E.P.G.V., F.V., S.v.S., N.M.W., Y.G.W., J.W., C.D.W., P.W-W., M.W., S.W., and Fr.Zi. Control cohorts: L.C.B., J.G.E., A.F., C.G., Ha.Ha., Y.-L.L., Wo.Li., J.L.M., A.M.M., M.M.N., Aa.Pa., F.P., K.S., H.S., G.N.T., and W.Y. Consortium coordination: K.E.B. and K.L.O.

# References


1. Thurman, D. J. et al. Standards for epidemiologic studies and surveillance of epilepsy. Epilepsia 52, 2–26 (2011).
2. Scheffer, I. E. et al. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. Epilepsia 58, 512–521 (2017).
3. Steinlein, O. K. et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat. Genet. 11, 201–203 (1995).
4. Helbig, I. et al. Primer Part 1 - the building blocks of epilepsy genetics. Epilepsia 57, 861–868 (2016).
5. McTague, A., & Howell, K. B., Cross, J. H., Kurian, M. A. & Scheffer, I. E. The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol. 15, 304–316 (2016).
6. Epi4K consortium and Epilepsy Phenome/Genome Project. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. Lancet Neurol. 16, 135–143 (2017).
7. Koeleman, B. P. C. What do genetic studies tell us about the heritable basis of common epilepsy? Polygenic or complex epilepsy? Neurosci. Lett. 667, 10–16 (2018).
8. Speed, D. et al. Describing the genetic architecture of epilepsy through heritability analysis. Brain 137, 2680–2689 (2014).
9. Vadlamudi, L. et al. Genetics of epilepsy: the testimony of twins in the molecular era. Neurology 83, 1042–1048 (2014).
10. Cavalleri, G. L. et al. Multicentre search for genetic susceptibility loci in sporadic epilepsy syndrome and seizure types: a case-control study. Lancet Neurol. 6, 970–980 (2007).
11. Kasperavičiūtė, D. et al. Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. Brain 133, 2136–2147 (2010).
12. EPICURE Consortium et al. Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. Hum. Mol. Genet. 21, 5359–5372 (2012).
13. Leu, C. et al. Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. Epilepsia 53, 308–318 (2012).
14. Guo, Y. et al. Two-stage genome-wide association study identifies variants in CAMSAP1L1 as susceptibility loci for epilepsy in Chinese. Hum. Mol. Genet 21, 1184–9 (2012).
15. ILAE Consortium on Complex Epilepsies. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. Lancet Neurol. 13, 893–903 (2014).
16. Oyrer, J. et al. Ion channels in genetic epilepsy: from genes and mechanisms to disease-targeted therapies. Pharmacol. Rev. 70, 142–173 (2018).

17. Okada, Y. et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506, 376–81 (2014).
18. Imamura, M. et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. Nat. Commun. 7, 10531 (2016).
19. Plecko, B. et al. Pyridoxine responsiveness in novel mutations of the PNPO gene. Neurology 82, 1425–1433 (2014).
20. Stockler, S. et al. Pyridoxine dependent epilepsy and antiquitin deficiency. Mol. Genet. Metab. 104, 48–60 (2011).
21. Snedeker, J. et al. Unique spatiotemporal requirements for intraflagellar transport genes during forebrain development. PLoS One 12, e0173258 (2017).
22. Stottmann, R. W., Tran, P. V., Turbe-Doan, A. & Beier, D. R. Ttc21b is required to restrict sonic hedgehog activity in the developing mouse forebrain. Dev. Biol. 335, 166–178 (2009).
23. Ng, B. et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat. Neurosci. 20, 1418–1426 (2017).
24. Schulz, H. et al. Genome-wide mapping of genetic determinants influencing DNA methylation and gene expression in human hippocampus. Nat. Commun. 8, 1511 (2017).
25. Freytag, S., Burgess, R., Oliver, K. L. & Bahlo, M. Brain-coX: investigating and visualising gene co-expression in seven human brain transcriptomic datasets. Genome Med. 9, 55 (2017).
26. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods 9, 215–216 (2012).
27. Sniekers, S. et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat. Genet. 49, 1107–1112 (2017).
28. Boyle, A. P. et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 22, 1790–7 (2012).
29. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat. Genet. 50, 621–629 (2018).
30. Holmes, M. D., Brown, M. & Tucker, D. M. Are 'generalized' seizures truly generalized? Evidence of localized mesial frontal and frontopolar discharges in absence. Epilepsia 45, 1568–1579 (2004).
31. Carney, P. W., Masterton, R. A. J., Flanagan, D., Berkovic, S. F. & Jackson, G. D. The frontal lobe in absence epilepsy: EEG-fMRI findings. Neurology 78, 1157–1165 (2012).
32. Chowdhury, F. A. et al. Impaired cognitive function in idiopathic generalized epilepsy and unaffected family members: an epilepsy endophenotype. Epilepsia 55, 835–840 (2014).
33. Koepp, M. J., Thomas, R. H., Wandschneider, B., Berkovic, S. F. & Schmidt, D. Concepts and controversies of juvenile myoclonic epilepsy: still an enigmatic epilepsy. Expert Rev. Neurother. 14, 819–831 (2014).

34. Curwood, E. K. et al. Abnormal cortical thickness connectivity persists in childhood absence epilepsy. Ann. Clin. Transl. Neurol. 2, 456–464 (2015).

35. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 8, 1826 (2017).

36. Allen, A. S. et al. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. Lancet Neurol. 16, 135–143 (2017).

37. Delahaye-Duriez, A. et al. Rare and common epilepsies converge on a shared gene regulatory network providing opportunities for novel antiepileptic drug discovery. Genome Biol. 17, 245 (2016).

38. So, H.-C. et al. Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. Nat. Neurosci. 20, 1342–1349 (2017).

39. Lennox, W. G. The heredity of epilepsy as told by relatives and twins. J. Am. Med. Assoc. 146, 529–536 (1951).

40. Thomas, R. H. & Berkovic, S. F. The hidden genetics of epilepsy—a clinically important new paradigm. Nat. Rev. Neurol. 10, 283–292 (2014).

41. Annegers, J. F., Hauser, W. A., Anderson, V. E. & Kurland, L. T. The risks of seizure disorders among relatives of patients with childhood onset epilepsy. Neurology 32, 174–9 (1982).

42. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. Am. J. Hum. Genet. 91, 1011–1021 (2012).

43. Speed, D. et al. Reevaluation of SNP heritability in complex human traits. Nat. Genet. 49, 986–992 (2017).

44. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 42, 565–569 (2010).

45. Japaridze, G. et al. Focal EEG features and therapeutic response in patients with juvenile absence and myoclonic epilepsy. Clin. Neurophysiol. 127,1182–1187 (2016).

46. Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat. Genet. 50, 229–237 (2018).

47. Wang, H.-G. et al. The auxiliary subunit KChIP2 is an essential regulator of homeostatic excitability. J. Biol. Chem. 288, 13258–13268 (2013).

48. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427 (2014).

49. Vlaskamp, D. R. M. et al. Haploinsufficiency of the STX1B gene is associated with myoclonic astatic epilepsy. Eur. J. Paediatr. Neurol. 20, 489–92 (2016).

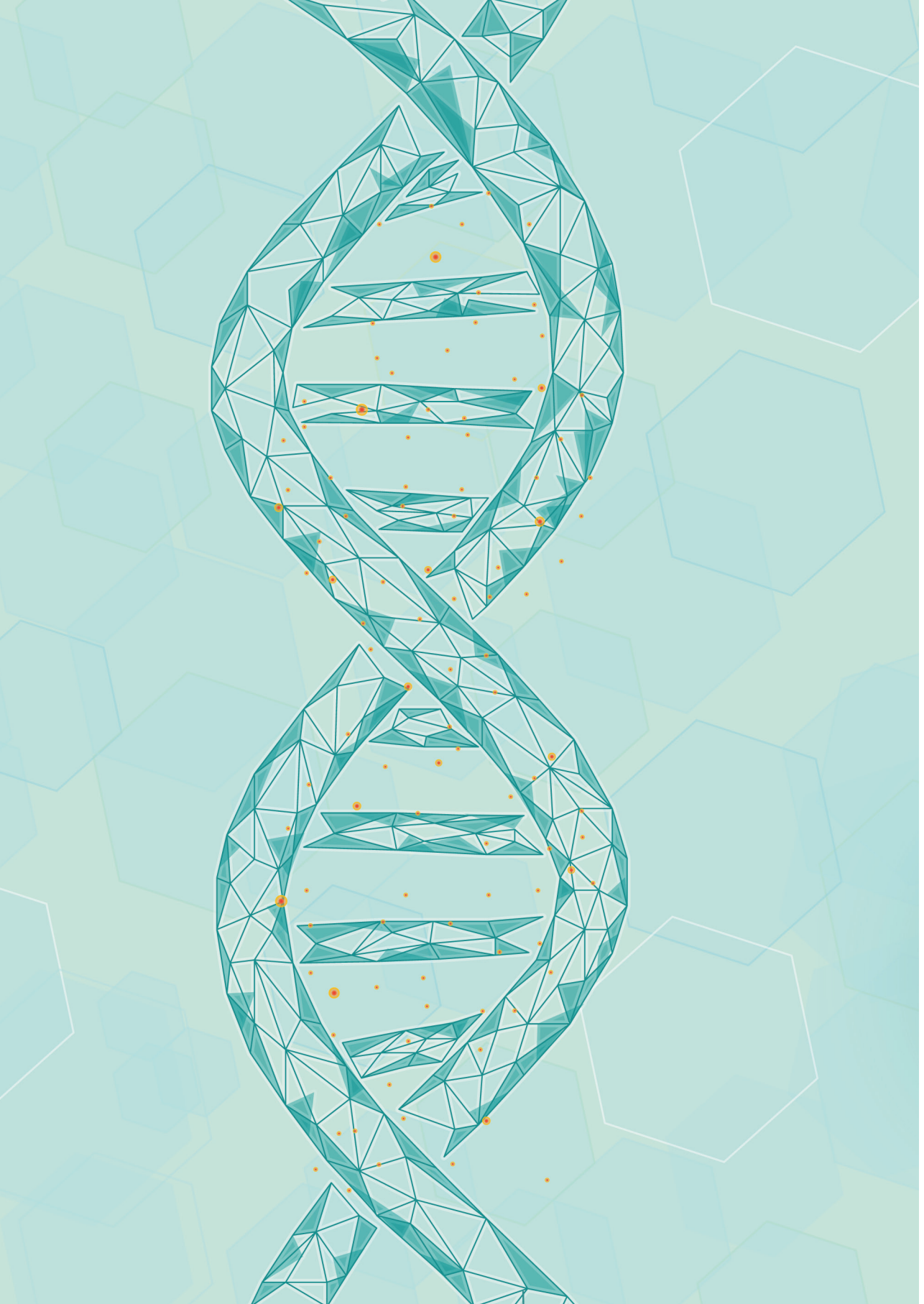
50. Schubert, J. et al. Mutations in STX1B, encoding a presynaptic protein, cause fever-associated epilepsy syndromes. Nat. Genet. 46, 1327–1332 (2014).

51. Wiencken-Barger, A. E., Djukic, B., Casper, K. B. & McCarthy, K. D. A role for Connexin43 during neurodevelopment. Glia 55, 675–86 (2007).

52. Collignon, F. et al. Altered expression of connexin subtypes in mesial temporal lobe epilepsy in humans. J. Neurosurg. 105, 77–87 (2006).

53. Garbelli, R. et al. Expression of connexin 43 in the human epileptic and drugresistant cerebral cortex. Neurology 76, 895–902 (2011).
54. Berg, A. T. et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. Epilepsia 51, 676–685 (2010).
55. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–75 (2007).
56. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat. Genet. 48, 1279–1283 (2016).
57. Menashe, I., Rosenberg, P. S. & Chen, B. E. PGA: power calculator for case-control genetic association analyses. BMC Genet. 9, 36 (2008).
58. Loh, P.-R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat. Genet. 47, 284–90 (2015).
59. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191 (2010).
60. de Bakker, P. I. W. et al. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum. Mol. Genet. 17, 122–128 (2008).
61. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. 47, 291–295 (2015).
62. Schmitt, A. D. et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep. 17, 2042–2059 (2016).
63. Fritsche, L. G. et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat. Genet. 48, 134–43 (2016).
64. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. 48, 245–252 (2016).
65. Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. 19, 1442–1453 (2016).
66. Aguet, F. et al. Genetic effects on gene expression across human tissues. Nature 550, 204–213 (2017).
67. McLaren, W. et al. The ensembl variant effect predictor. Genome Biol. 17, 122 (2016).
68. Rossin, E. J. et al. Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS Genet. 7, e1001273 (2011).
69. Lage, K. et al. A human phenome-interactome network of protein complexes implicated in genetic disorders. Nat. Biotechnol. 25, 309–316 (2007).
70. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015).
71. ENCODE Project Consortium, I. et al. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).

72. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 47, 1228–35 (2015).
73. Santos, R. et al. A comprehensive map of molecular drug targets. Nat. Rev. Drug Discov. 16, 19–34 (2017).
74. Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell 171, 1437–1452 (2017). e17.
75. Cheng, J. et al. Evaluation of analytical methods for connectivity map data. Pac. Symp. Biocomput. 2013, 5–16 (2013).
76. Duan, Q. et al. L1000CDS2: LINCS L1000 characteristic direction signatures search engine. NPJ Syst. Biol. Appl. 2, 16015 (2016).
77. Sivapalarajah, S. et al. The prescribable drugs with efficacy in experimental epilepsies (PDE3) database for drug repurposing research in epilepsy. Epilepsia 59, 492–501 (2018).
78. Semah, F. et al. Is the underlying cause of epilepsy a major prognostic factor for recurrence? Neurology 51, 1256–1262 (1998).
79. Jallon, P. & Latour, P. Epidemiology of idiopathic generalized epilepsies. Epilepsia 46(Suppl 9), 10–4 (2005).
80. Loh, P.-R. et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. Nat. Genet. 47, 1385–92 (2015).
81. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 33, 272–279 (2017).



## CHAPTER 3

# ASSESSING THE GENETIC ASSOCIATION BETWEEN VITAMIN B6 METABOLISM AND GENETIC GENERALIZED EPILEPSY

Remi Stevelink, Faith Pangilinan, Floor E. Jansen, Kees P.J. Braun, International League Against Epilepsy Consortium on Complex Epilepsies, Anne M. Molloy, Lawrence C. Brody, Bobby P.C. Koeleman

*Molecular Genetics and Metabolism Reports.* 2019; Oct 11;21:100518.

## Abstract

Altered vitamin B6 metabolism due to pathogenic variants in the gene PNPO causes early onset epileptic encephalopathy, which can be treated with high doses of vitamin B6. We recently reported that single nucleotide polymorphisms (SNPs) that influence PNPO expression in the brain are associated with genetic generalized epilepsy (GGE). However, it is not known whether any of these GGE-associated SNPs influence vitamin B6 metabolite levels. Such an influence would suggest that vitamin B6 could play a role in GGE therapy. Here, we performed genome-wide association studies (GWAS) to assess the influence of GGE associated genetic variants on measures of vitamin B6 metabolism in blood plasma in 2232 healthy individuals. We also asked if SNPs that influence vitamin B6 were associated with GGE in 3122 affected individuals and 20,244 controls. Our GWAS of vitamin B6 metabolites reproduced a previous association and found a novel genome-wide significant locus. The SNPs in these loci were not associated with GGE. We found that 84 GGE-associated SNPs influence expression levels of PNPO in the brain as well as in blood. However, these SNPs were not associated with vitamin B6 metabolism in plasma. By leveraging polygenic risk scoring (PRS), we found suggestive evidence of higher catabolism and lower levels of the active and transport forms of vitamin B6 in GGE, although these findings require further replication.

# Introduction

Treatment with vitamin B6 can control seizures in a subset of children with early-onset intractable seizures [1]. Such vitamin B6-responsive epilepsy can be caused by mutations in a number of genes, particularly pyridoxal-5′-phosphate oxidase (PNPO [[2], [3], [4], [5]]), which is essential to convert pyridox(am)ine-5′-phosphate into the active form of vitamin B6, pyridoxal-5′-phosphate (PLP). In mammals, PLP is a cofactor for >160 different enzymatic reactions, including the metabolism of the neurotransmitters glutamate and GABA [6].

Interestingly, PLP levels are also reduced in some patients with common forms of epilepsy [7,8], possibly due to the effects of anti-epileptic drugs [9,10]. Moreover, dietary depletion of PLP can induce seizures and epileptiform EEG abnormalities in healthy individuals [11,12]. PLP treatment can reduce seizure frequency in some refractory epilepsy patients without documented pathogenic variants [7,13].

Our recent genome-wide association study (GWAS) of genetic generalized epilepsy (GGE) confirmed and strengthened a genome-wide significant association between GGE and a haplotype containing PNPO as the most likely causal gene [14]. We found that GGE-associated SNPs alter expression of PNPO in the dorsolateral prefrontal cortex [14], suggesting that altered vitamin B6 metabolism might be involved in the pathophysiology of GGE. If so, metabolic pathways involving vitamin B6 might be a therapeutic target. However, it is unknown whether SNPs that influence metabolite levels in blood also predispose to GGE. Likewise, it is not known if GGE associated SNPs are associated with changes in vitamin B6 levels in blood. We sought to answer both of these related questions.

Here, we assessed the genetic association of vitamin B6 metabolites with GGE, utilizing data from two independent large studies, one which compared genetic variants between people with and without epilepsy [14] and the other which evaluated genetic influences of blood vitamin B6 metabolites in healthy individuals [15]. Our previously reported GWAS on 2232 healthy individuals assessed the influence of genetic variants on three different pyridoxine metabolite concentrations measured in blood: PLP, the cell-

membrane transport form pyridoxal (PL), and the catabolite pyridoxic acid (PA) [15]. To fully capture genetic contribution to vitamin B6 metabolism, we repeated these genome-wide analyses with imputed genotypes and examined two additional, derived markers of pyridoxine metabolism [6]: the ratios (“PLP:PL”) and  (“PAr index”). We then used two approaches to determine whether genetic contribution to vitamin B6 metabolism or GGE might be reciprocally informative. First, we assessed whether the GGE-associated SNPs that alter PNPO expression are associated with these 5 measures of pyridoxine metabolism. Second, we utilized polygenic risk scoring (PRS) to assess whether the SNPs that influence pyridoxine metabolism are also associated with GGE by comparing PRS for the five measures of pyroxidine metabolism between 3122 people with GGE and 20,244 controls.

# Methods

## Subjects

A sample of 2232 healthy individuals from the Trinity Student Study (TSS) were studied to assess genetic variants that influence pyridoxine metabolism. TSS participants are ethnically Irish people aged 18 to 28 years without any serious medical conditions [15,16].

A subset of 3122 non-related subjects with GGE and 20,244 controls from the epilepsy GWAS of the ILAE Consortium on Complex Epilepsies were studied for PRS analyses [14]. These consisted of a subset of the subjects with European ancestry drawn from the more ethnically diverse subjects in the original GWAS. Moreover, the TSS, which served as a control cohort for the original epilepsy GWAS, was excluded from the current PRS analyses. Approval was obtained by all relevant institutional review boards and all study participants provided written informed consent.

## Measurement of pyridoxine metabolism

We collected non-fasting EDTA blood samples for measurement of B6 vitamer concentrations in the TSS cohort. Samples were centrifuged and plasma was frozen within 3h of phlebotomy. Details of stability of B6 vitamers under the conditions of collection have been determined [17]. B6 vitamers were

measured using liquid chromatography–tandem mass spectrometry, as described previously [15]. The methodology included measurements of the primary B6 vitamers (PLP, PL and PA) plus the less abundant vitamers (pyridoxamine, pyridoxamine phosphate, pyridoxine, pyridoxine phosphate). These latter vitamers were below the limit of detection in most samples and were not included in the GWAS analysis. From the primary B6 vitamers, the ratios ("PLP:PL") and ("PAr index") were calculated. The amount of vitamin B6 intake from supplements and fortified foods was quantified using a standardized questionnaire in which study participants reported their recent intake from a list of commonly used vitamin supplements. Active nutrient information was obtained for each supplement and converted to μg of nutrient per day as previously described [15,16].

## Genotyping and quality control

We performed genotyping, imputation and genotype quality control for the TSS and the ILAE cohorts identically for both cohorts, as described earlier [14]. In addition, we excluded related subjects from each cohort. Genetic relatedness was calculated in the TSS with PLINK [18,19] and in the ILEA using KING [20] and one individual from each pair with 3rd degree or stronger relatedness (kinship coefficient > 0.0442) was retained.

## Pyridoxine metabolite GWAS

We repeated the previously published GWAS on log-transformed PLP, PL and PA levels [15]. In addition, we now used imputed genotype data (~6 million SNPs instead of ~750 thousand) and quality control procedures that were the same for TSS and the epilepsy GWAS. We also performed a GWAS on PLP:PL and the PAr index. We performed linear-mixed model association analyses with Emmax [21], and included age, gender and log-transformed vitamin B6 supplement intake as covariates. Genome-wide significance was defined as P 5 × 10−8.

## PNPO eQTL analyses

Summary statistics from the previously published GGE GWAS [14] were used to specifically assess SNPs in the previously found genome-wide significant

PNPO locus. This locus was defined as the lead SNP (SNP with the lowest P-value) and all SNPs that were in linkage disequilibrium with the lead SNP ($R2 > 0.2$) and had a P-value$<10^{-4}$. We next used FUMA [22] to assess which of these variants is significantly associated with PNPO expression in blood, using expression quantitative trait loci (eQTL) data from the eQTLgen study (based on RNA-sequencing data from 24,886 whole blood and n = 4798 peripheral blood mononuclear cell samples [23]). Finally, we assessed the association P-value of these SNPs in the 5 different pyridoxine metabolism GWAS.

### Polygenic risk score analyses

We used default settings of PRSice to perform PRS analyses to establish whether people with GGE have different pyridoxine metabolism PRS scores compared to controls. In brief, every SNP was assigned a weight according to its association in the 5 different pyridoxine metabolism GWAS. Individual PRS were than calculated as the sum of weighted effect alleles, standardized using a Z-transformation: . Only high-quality SNPs with a genotype call-rate > 0.99 and a minor allele frequency > 0.01 were used. SNPs were pruned to a subset of uncorrelated SNPs ($R2 < 0.1$) and PRS values were calculated with a range of different P-value thresholds from 0.0001 to 0.5, in steps of 0.0005 (default for PRSice). Logistic regression analyses were used to assess whether pyridoxine metabolite PRS scores were significantly different in people with GGE compared to controls, while controlling for 10 principal components of ancestry. The 'best-fit' P-value threshold was selected, defined as the PRS with the strongest association with GGE. We corrected for multiple testing by using a conservative significance threshold of $P < .001$, as recommended for PRSice [24]. We calculated the explained variance (Nagelkerke's R2) by subtracting the full logistic regression model (PRS + covariates) with the null model (covariates only).

## Results

### Genetic variants that influence vitamin B6 metabolite levels

Our GWAS analyses on vitamin B6 measures (Fig. 1) replicated the genome-wide significant signal at 1p36.12, implicating the ALPL gene (Fig. 1A).

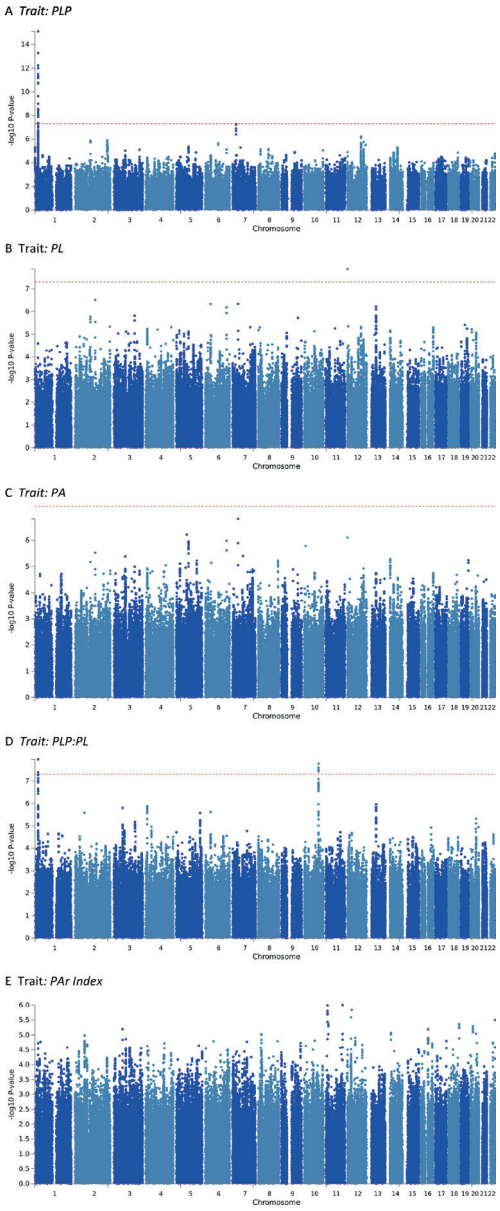



**Figure 1: Manhattan plots for each genome-wide association analysis of the five measures of vitamin B6 metabolism.** Each genome-wide association analysis was performed using an imputed SNP set and log10-transformed values. **A**) pyridoxal 5′-phosphate (PLP), **B**) pyridoxal (PL), **C**) pyridoxic acid (PA), **D**) PLP:PL ratio, **E**) PAr index. X-axis: Tested SNPs according to chromosomal position. Y-axis: Negative log10-transformed p-values. Red line: Genome-wide significance ($p < 5 \times 10-8$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This signal was significantly associated with PLP concentrations as well as the PLP:PL ratio (p = 7.4*10−16 and p = 1.1*10−8, respectively). In addition, we found a novel PLP:PL locus at 10q24.2 (Fig. 1D), which includes a missense variant of the gene pyridine nucleotide disulfide oxidoreductase domain 2 (PYROXD2, rs2147896; p = 3.7*10−8; Fig. 2). We note an additional locus associated with PLP (Fig. 1A) in an intergenic region on chromosome 7 that is just under the threshold for genome-wide significance (lead SNP rs61295180, p = 5.6*10−8). Last, there is a single imputed SNP on chromosome 12 associated with PL (rs4765900, Fig. 1B), but there is no LD signature and this singleton is likely to be a spurious signal.

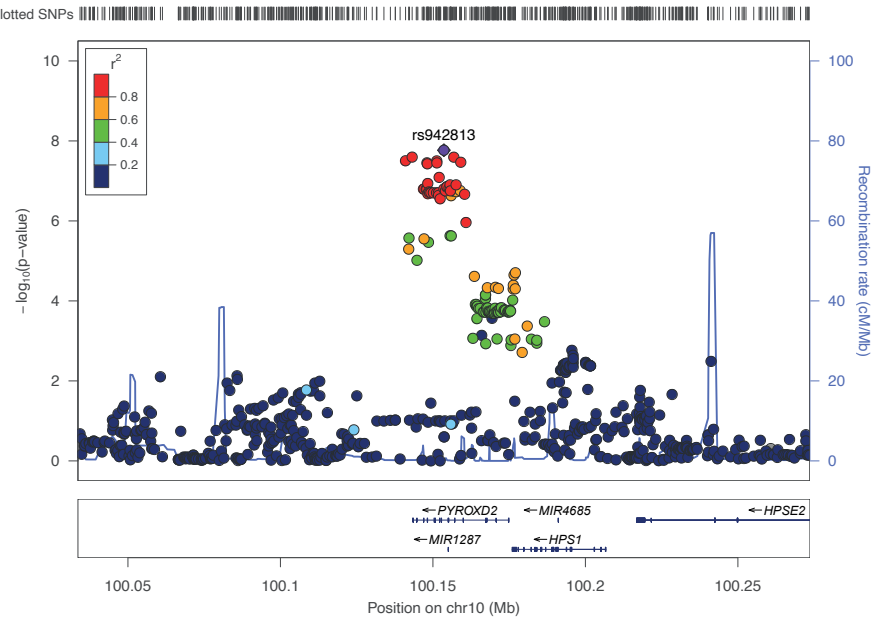



**Figure 2: Locusplot of the genome-wide significant locus associated with the PLP:PL ratio**. This locus includes a missense variant of the gene PYROXD2 (rs2147896). Chromosomal position including gene annotations are displayed on the X-axis and negative log10-transformed P-values are displayed on the Y-axis. SNPs are plotted as circles whose colors represent the correlation (linkage disequilibrium) with the lead SNP rs942813.

## Effect of PNPO SNPs on vitamin B6 metabolite levels in blood

A genome-wide significant locus identified in the GGE GWAS implicated 84 SNPs around the gene PNPO [14]. To assess whether these SNPs influence

vitamin B6 levels in blood, we leveraged an eQTL database including data from 29,684 subjects and found that all 84 SNPs are associated with PNPO expression in blood (eQTL p-values between 1.6*10−85 and 8.5*10−8; see Supplementary Table 1).

We tested these 84 SNPs for association with vitamin B6 metabolites levels and found that only 20 of these SNPs reached nominal significance ($p < .05$) for association with any of the 5 measures of vitamin B6 metabolism (Supplementary Table 2). None survived correction for multiple comparisons, suggesting that GGE-associated variants that influence PNPO expression are not associated with concentrations of vitamin B6 metabolites in blood.



## Polygenic association of vitamin B6 metabolite SNPs with GGE

To assess whether vitamin B6 metabolism is different in GGE, we first assessed whether the genome-wide significant SNPs from the pyridoxine metabolite GWAS showed an association with GGE. Two out of 44 SNPs from the ALPL locus and none from the PYROXD2 locus from the PLP and PLP:PL GWAS showed a nominally significant association with GGE, but these did not survive correction for multiple testing (Supplementary Table 3).

At just over 2000 individuals, the vitamin B6 metabolite GWAS had limited power to detect associations at the stringent genome-wide significance threshold. It is likely that there are additional, undetected genetic variants with smaller effect sizes than we had the power to detect that influence vitamin B6 metabolite levels. Therefore, we next used PRS analyses to leverage the full distribution of SNPs from the pyridoxine metabolite GWAS, to assess whether people with GGE have a genetic predisposition for different metabolite levels compared to controls. Briefly, polygenic risk scores were generated by determining which genomic SNPs collectively contribute to vitamin B6 metabolite measures in the TSS above a set threshold. These SNPs were used to generate PRS scores in GGE participants and controls to ask whether genetic contribution to vitamin B6 metabolism differs in these groups.

These polygenic score analyses showed a trend towards lower scores for PLP and PL, but higher scores for PA, PLP:PL and the PAr index in GGE participants (Fig. 3; see Supplementary Table 4 for values). However,

these associations did not meet the stringent P < .001 threshold that is recommended for analyses with PRSice [24].

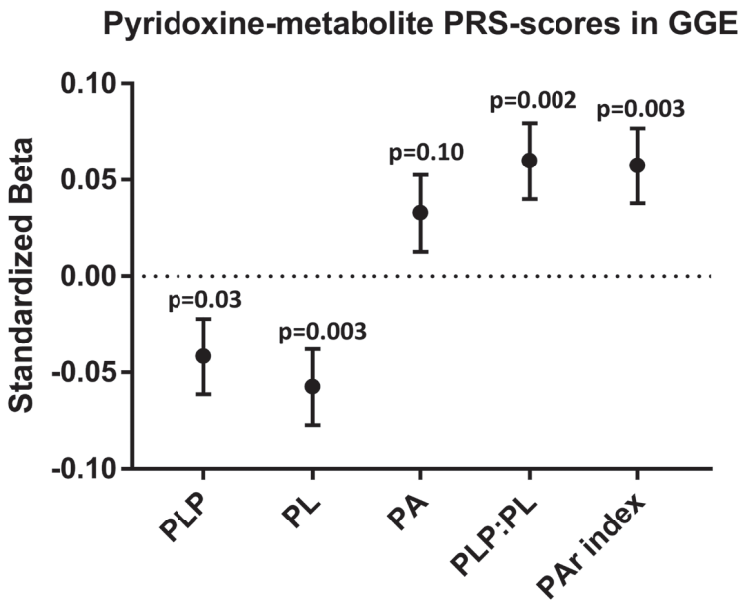



**Figure 3: Logistic regression to assess the difference in pyridoxine-related metabolite PRS scores between people with GGE compared to controls.** Standardized beta regression coefficients ±standard error are displayed. See Supplementary Table 2 for values. None of the associations reached the significance threshold of P < .001 that is recommended for analyses with PRSice.

# Discussion

In this study, we assessed the genetic association between vitamin B6 metabolism and GGE. We previously found that all 84 GGE-associated SNPs in the PNPO locus significantly influenced gene expression of PNPO, which is essential to convert vitamin B6 into its active form PLP. In this study, these SNPs were not associated with alterations in vitamin B6 metabolite levels or ratios in blood plasma. However, we cannot rule out the possibility that these SNPs influence vitamin B6 metabolism in the brain or other tissues, where it is needed to convert pyridox(am)ine 5-phosphate to the active form PLP. Indeed, the correlation of PLP as measured in CSF or plasma in children with an intellectual disability was found to be significant but not complete

[9], indicating that genetic influence on PLP and the ability to detect it may differ in CSF and plasma. It is possible that GGE-associated PNPO SNPs specifically influence metabolism of vitamin B6 in the brain, which could affect neurotransmitter metabolism and influence seizure susceptibility, without altering detectable levels in plasma. However, it is not feasible to collect CSF samples at the scale required for GWAS analyses to ask this more directly. Another caveat to consider is the potentially reductive effect anti-epileptic drugs may have on PLP in the GGE population. This potential influence may have reduced our ability to detect contribution of genetic modifiers of vitamin B6 metabolism to epilepsy in the GGE population.

We reproduced previous GWAS of vitamin B6 levels [15], which confirmed the locus around the gene ALPL, which codes for the enzyme that converts PLP into PL. In addition, by performing a GWAS on PLP:PL, we found a new locus implicating the gene PYROXD2. The same SNPs in this locus are associated with levels of several other metabolites in blood (dimethylamine [25], unknown X-12092 [26,27], caprolactam [28], asymmetric dimethylarginine [29]) and urine (trimethylamine [25,30]) but a role for PYROXD2 in vitamin B6 metabolism has not been previously established. PYROXD2 was initially identified for binding the X protein of human hepatitis B (HBx) in a protein interaction assay [31]. It has been further characterized as having tumor suppressor activity [32], although its exact function remains unknown. Its protein sequence includes sequence conservation with an NAD(P)-binding Rossman-like domain (HomoloGene [33]), which may contribute to a reduction-oxidation activity. The association of PYROXD2 with a measure of vitamin B6 metabolism in the current study may be a helpful clue in elucidating its biological function.

Although we found that the genome-wide significant vitamin B6 metabolism loci were not significantly associated with GGE, we did find suggestive evidence for an association by leveraging the full distribution of SNPs with PRS analyses. These analyses suggested that people with GGE have a genetic predisposition for higher vitamin B6 catabolism (higher PA and PAr index) and lower levels of PLP and PL in blood. Moreover, PRS of PLP:PL was higher in GGE compared to controls, suggesting relatively lower levels of the transport form PL, which is required for delivery to the brain. However,

these analyses were limited by a relatively small sample size for the vitamin B6 GWAS (n = 2232) and did not meet the stringent $P < .001$ cutoff that is recommended for PRSice. Further studies with a larger sample size are needed to confirm these findings.

In summary, our study did not find evidence for an influence of GGE-associated PNPO SNPs on vitamin B6 metabolism in blood, although these SNPs could still have a brain-specific influence on vitamin B6. We found a novel locus that influences the PLP:PL ratio and we found suggestive evidence for increased vitamin B6 catabolism in people with GGE, which needs further replication. However, it is unlikely that genetic differences in vitamin B6 metabolism described here are sufficiently large to be causal in the pathophysiology of GGE or to have direct therapeutic implications.

## Acknowledgements

We are grateful to the Ming Fund for supporting this project. This work was also supported by the Intramural Research Program of the National Human Genome Research Institute. Parts of the analysis of this work were performed on resources of the High Performance Center of the University of Luxembourg and Elixir-Luxembourg. Some of the data reported in this paper were collected as part of a project undertaken by the International League against Epilepsy (ILAE) and some of the authors are experts selected by the ILAE. Opinions expressed by the authors, however, do not necessarily represent the policy or position of the ILAE.

## Declaration of Competing Interest

None of the authors has any conflict of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

## Supporting information

Additional supporting information may be found at: https://tinyurl.com/4j3srm5j.

# References


1. R. Ramachandrannair, M. Parameswaran, Prevalence of pyridoxine dependent seizures in south Indian children with early onset intractable epilepsy: a hospital based prospective study, Eur. J. Paediatr. Neurol. 9 (2005) 409–413.
2. B. Plecko, K. Paul, P. Mills, P. Clayton, E. Paschke, O. Maier, O. Hasselmann, G. Schmiedel, S. Kanz, M. Connolly, N. Wolf, E. Struys, S. Stockler, L. Abela, D. Hofer, Pyridoxine responsiveness in novel mutations of the PNPO gene, Neurology 82 (2014) 1425–1433.
3. B. Jaeger, N.G. Abeling, G.S. Salomons, E.A. Struys, M. Simas-Mendes, V.G. Geukers, B.T. Poll-The, Pyridoxine responsive epilepsy caused by a novel homozygous PNPO mutation, Mol. Genet. Metab. Rep. 6 (2016) 60–63.
4. P.B. Mills, S.S. Camuzeaux, E.J. Footitt, K.A. Mills, P. Gissen, L. Fisher, K.B. Das, S.M. Varadkar, S. Zuberi, R. McWilliam, T. Stodberg, B. Plecko, M.R. Baumgartner, O. Maier, S. Calvert, K. Riney, N.I. Wolf, J.H. Livingston, P. Bala, C.F. Morel, F. Feillet, F. Raimondi, E. Del Giudice, W.K. Chong, M. Pitt, P.T. Clayton, Epilepsy due to PNPO mutations: genotype, environment and treatment affect presentation and outcome, Brain 137 (2014) 1350–1360.
5. T.L. Ware, J. Pitt, J. Freeman, Pyridoxine responsiveness in novel mutations of the PNPO gene, Neurology 84 (2015) 329.
6. P.M. Ueland, A. Ulvik, L. Rios-Avila, O. Midttun, J.F. Gregory, Direct and functional biomarkers of vitamin B6 status, Annu. Rev. Nutr. 35 (2015) 33–70.
7. M. Goyal, P.R. Fequiere, T.M. McGrath, K. Hyland, Seizures with decreased levels of pyridoxal phosphate in cerebrospinal fluid, Pediatr. Neurol. 48 (2013) 227–231.
8. E.J. Footitt, S.J. Heales, P.B. Mills, G.F. Allen, M. Oppenheim, P.T. Clayton, Pyridoxal 5′-phosphate in cerebrospinal fluid; factors affecting concentration, J. Inherit. Metab. Dis. 34 (2011) 529–538.
9. M. Albersen, M. Bosma, J.J. Jans, F.C. Hofstede, P.M. van Hasselt, M.G. de Sain-van der Velden, G. Visser, N.M. Verhoeven-Duif, Vitamin B6 in plasma and cerebrospinal fluid of children, PLoS ONE 10 (2015) e0120872.
10. T. Apeland, M.A. Mansoor, K. Pentieva, H. McNulty, R.E. Strandjord, Fasting and post-methionine loading concentrations of homocysteine, vitamin B2, and vitamin B6 in patients on antiepileptic drugs, Clin. Chem. 49 (2003) 1005–1008.
11. D.B. Coursin, Convulsive seizures in infants with pyridoxine-deficient diet, J. Am. Med. Assoc. 154 (1954) 406–408.
12. M.J. Kretsch, H.E. Sauberlich, E. Newbrun, Electroencephalographic changes and periodontal status during short-term vitamin B-6 depletion of young, nonpregnant women, Am. J. Clin. Nutr. 53 (1991) 1266–1274.
13. H.S. Wang, M.F. Kuo, M.L. Chou, P.C. Hung, K.L. Lin, M.Y. Hsieh, M.Y. Chang, Pyridoxal phosphate is better than pyridoxine for controlling idiopathic intractable epilepsy, Arch. Dis. Child. 90 (2005) 512–515.
14. E. International League, Against epilepsy consortium on complex, genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies, Nat. Commun. 9 (2018) 5269.

15. T.C. Carter, F. Pangilinan, A.M. Molloy, R. Fan, Y. Wang, B. Shane, E.R. Gibney, O. Midttun, P.M. Ueland, C.D. Cropp, Y. Kim, A.F. Wilson, J.E. Bailey-Wilson, L.C. Brody, J.L. Mills, Common variants at putative regulatory sites of the tissue nonspecific alkaline phosphatase gene influence circulating pyridoxal 5'-phosphate concentration in healthy adults, J. Nutr. 145 (2015) 1386–1393.
16. J.L. Mills, T.C. Carter, J.M. Scott, J.F. Troendle, E.R. Gibney, B. Shane, P.N. Kirke, P.M. Ueland, L.C. Brody, A.M. Molloy, Do high blood folate concentrations exacerbate metabolic abnormalities in people with low vitamin B-12 status? Am. J. Clin. Nutr. 94 (2011) 495–500.
17. O. Midttun, M.K. Townsend, O. Nygard, S.S. Tworoger, P. Brennan, M. Johansson, P.M. Ueland, Most blood biomarkers related to vitamin status, one-carbon metabolism, and the kynurenine pathway show adequate preanalytical stability and within-person reproducibility to allow assessment of exposure or nutritional status in healthy women and cardiovascular patients, J. Nutr. 144 (2014) 784–790.
18. A.M. Molloy, F. Pangilinan, J.L. Mills, B. Shane, M.B. O'Neill, D.M. McGaughey, A. Velkova, H.O. Abaan, P.M. Ueland, H. McNulty, M. Ward, J.J. Strain, C. Cunningham, M. Casey, C.D. Cropp, Y. Kim, J.E. Bailey-Wilson, A.F. Wilson, L.C. Brody, A common polymorphism in HIBCH influences methylmalonic acid concentrations in blood independently of cobalamin, Am. J. Hum. Genet. 98 (2016) 869–882.
19. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. de Bakker, M.J. Daly, P.C. Sham, PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (2007) 559–575.
20. A. Manichaikul, J.C. Mychaleckyj, S.S. Rich, K. Daly, M. Sale, W.M. Chen, Robust relationship inference in genome-wide association studies, Bioinformatics 26 (2010) 2867–2873.
21. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies, Nat. Genet. 44 (2012) 821–824.
22. K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma, Functional mapping and annotation of genetic 9, 10 associations with FUMA, Nat. Commun. 8 (2017) 1826.
23. U. Võsa, A. Claringbould, H.-J. Westra, M.J. Bonder, P. Deelen, B. Zeng, H. Kirsten, A. Saha, R. Kreuzhuber, S. Kasela, N. Pervjakova, I. Alvaes, M.-J. Fave, M. Agbessi, M. Christiansen, R. Jansen, I. Seppälä, L. Tong, A. Teumer, K. Schramm, G. Hemani, J. Verlouw, H. Yaghootkar, R. Sönmez, A.A. Andrew, V. Kukushkina, A. Kalnapenkis, S. Rüeger, E. Porcu, J. Kronberg-Guzman, J. Kettunen, J. Powell, B. Lee, F. Zhang, W. Arindrarto, F. Beutner, H. Brugge, J. Dmitrieva, M. Elansary, B.P. Fairfax, M. Georges, B.T. Heijmans, M. Kähönen, Y. Kim, J.C. Knight, P. Kovacs, K. Krohn, S. Li, M. Loeffler, U.M. Marigorta, H. Mei, Y. Momozawa, M. Müller- Nurasyid, M. Nauck, M. Nivard, B. Penninx, J. Pritchard, O. Raitakari, O. Rotzschke, E.P. Slagboom, C.D.A. Stehouwer, M. Stumvoll, P. Sullivan, P.A.C. 't Hoen, J. Thiery, A. Tönjes, J. van Dongen, M. van Iterson, J. Veldink, U. Völker, C. Wijmenga, M. Swertz, A. Andiappan, G.W. Montgomery, S. Ripatti, M. Perola, Z. Kutalik, E. Dermitzakis, S. Bergmann, T. Frayling, J. van Meurs, H. Prokisch, H. Ahsan, B. Pierce, T. Lehtimäki, D. Boomsma, B.M. Psaty, S.A. Gharib, P. Awadalla, L. Milani, W.H. Ouwehand, K. Downes, O. Stegle, A. Battle, J. Yang, P.M. Visscher, M. Scholz, G. Gibson, T. Esko, L. Franke, Unraveling the Polygenic Architecture of Complex Traits Using Blood eQTL Meta-Analysis bioRxiv, (2018).

24. J. Euesden, C.M. Lewis, P.F. O'Reilly, PRSice: polygenic risk score software, Bioinformatics 31 (2015) 1466–1468.

25. G. Nicholson, M. Rantalainen, J.V. Li, A.D. Maher, D. Malmodin, K.R. Ahmadi, J.H. Faber, A. Barrett, J.L. Min, N.W. Rayner, H. Toft, M. Krestyaninova, J. Viksna, S.G. Neogi, M.E. Dumas, U. Sarkans, P.C. Mol, P. Donnelly, T. Illig, J. Adamski, K. Suhre, M. Allen, K.T. Zondervan, T.D. Spector, J.K. Nicholson, J.C. Lindon, D. Baunsgaard, E. Holmes, M.I. McCarthy, C.C. Holmes, A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection, PLoS Genet. 7 (2011) e1002270.

26. S.Y. Shin, E.B. Fauman, A.K. Petersen, J. Krumsiek, R. Santos, J. Huang, M. Arnold, I. Erte, V. Forgetta, T.P. Yang, K. Walter, C. Menni, L. Chen, L. Vasquez, A.M. Valdes, C.L. Hyde, V. Wang, D. Ziemek, P. Roberts, L. Xi, E. Grundberg, C. Multiple Tissue Human Expression Resource, M. Waldenberger, J.B. Richards, R.P. Mohney, M.V. Milburn, S.L. John, J. Trimmer, F.J. Theis, J.P. Overington, K. Suhre, M.J. Brosnan, C. Gieger, G. Kastenmuller, T.D. Spector, N. Soranzo, An atlas of genetic influences on human blood metabolites, Nat. Genet. 46 (2014) 543–550.

27. J. Krumsiek, K. Suhre, A.M. Evans, M.W. Mitchell, R.P. Mohney, M.V. Milburn, B. Wagele, W. Romisch-Margl, T. Illig, J. Adamski, C. Gieger, F.J. Theis, G. Kastenmuller, Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information, PLoS Genet. 8 (2012) e1003005.

28. M.G. Hong, R. Karlsson, P.K. Magnusson, M.R. Lewis, W. Isaacs, L.S. Zheng, J. Xu, H. Gronberg, E. Ingelsson, Y. Pawitan, C. Broeckling, J.E. Prenni, F. Wiklund, J.A. Prince, A genome-wide assessment of variability in human serum metabolism, Hum. Mutat. 34 (2013) 515–524.

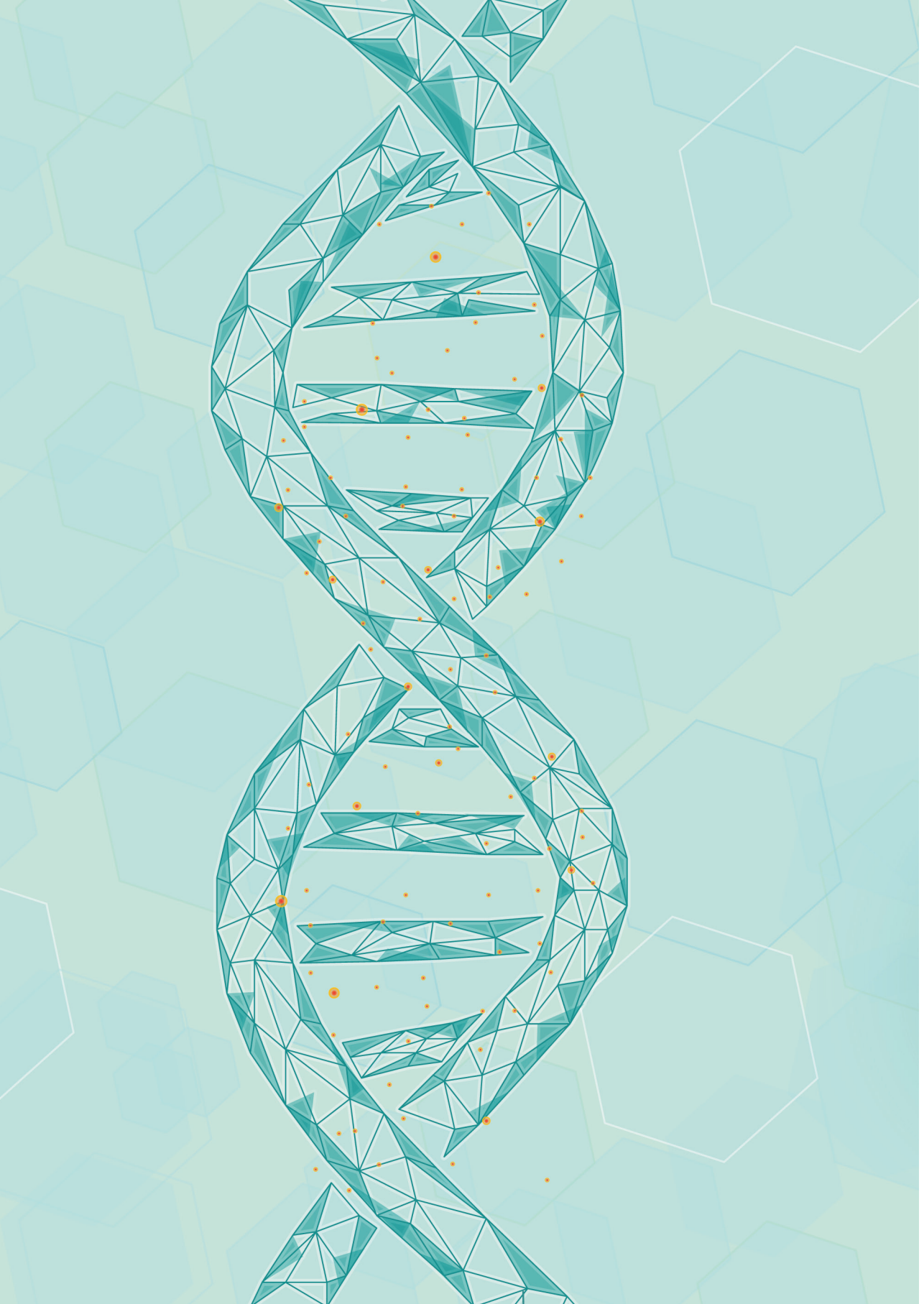
29. E.P. Rhee, J.E. Ho, M.H. Chen, D. Shen, S. Cheng, M.G. Larson, A. Ghorbani, X. Shi, I.T. Helenius, C.J. O'Donnell, A.L. Souza, A. Deik, K.A. Pierce, K. Bullock, G.A. Walford, R.S. Vasan, J.C. Florez, C. Clish, J.R. Yeh, T.J. Wang, R.E. Gerszten, A genome-wide association study of the human metabolome in a community-based cohort, Cell Metab. 18 (2013) 130–143.

30. R. Rueedi, M. Ledda, A.W. Nicholls, R.M. Salek, P. Marques-Vidal, E. Morya, K. Sameshima, I. Montoliu, L. da Silva, S. Collino, F.P. Martin, S. Rezzi, C. Steinbeck, D.M. Waterworth, G. Waeber, P. Vollenweider, J.S. Beckmann, J. Le Coutre, V. Mooser, S. Bergmann, U.K. Genick, Z. Kutalik, Genome-wide association study of metabolic traits reveals novel gene-metabolite-disease links, PLoS Genet. 10 (2014) e1004132.

31. J.L. Zhang, W.G. Zhao, K.L. Wu, K. Wang, X. Zhang, C.F. Gu, Y. Li, Y. Zhu, J.G. Wu, Human hepatitis B virus X protein promotes cell proliferation and inhibits cell apoptosis through interacting with a serine protease, Hepsin. Arch. Virol. 150 (2005) 721–741.

32. J. Huang, K. Wu, J. Zhang, W. Si, Y. Zhu, J. Wu, Putative tumor suppressor YueF affects the functions of hepatitis B virus X protein in hepatoma cell apoptosis and p53 expression, Biotechnol. Lett. 30 (2008) 235–242.

33. D.L. Wheeler, T. Barrett, D.A. Benson, S.H. Bryant, K. Canese, D.M. Church, M. DiCuccio, R. Edgar, S. Federhen, W. Helmberg, D.L. Kenton, O. Khovayko, D.J. Lipman, T.L. Madden, D.R. Maglott, J. Ostell, J.U. Pontius, K.D. Pruitt, G.D. Schuler, L.M. Schriml, E. Sequeira, S.T. Sherry, K. Sirotkin, G. Starchenko, T.O. Suzek, R. Tatusov, T.A. Tatusova, L. Wagner, E. Yaschenko, Database resources of the national center for biotechnology information, Nucleic Acids Res. 33 (2005) D39–D45.



## CHAPTER 4

# SHARED GENETIC BASIS BETWEEN GENETIC GENERALIZED EPILEPSY AND BACKGROUND ELECTROENCEPHALOGRAPHIC OSCILLATIONS

Remi Stevelink, Jurjen J. Luykx, Bochao D. Lin, Costin Leu, Dennis Lal, Alexander W. Smith, Dick Schijven, Johannes A. Carpay, Koen Rademaker, Roiza A. Rodrigues Baldez, Orrin Devinsky, Kees P. J. Braun, Floor E. Jansen, Dirk J. A. Smit, Bobby P. C. Koeleman, International League Against Epilepsy Consortium on Complex Epilepsies, Epi25 Collaborative

*Epilepsia*. 2021; Jul;62(7):1518-1527.

# Abstract

**Objective**

Paroxysmal epileptiform abnormalities on electroencephalography (EEG) are the hallmark of epilepsies, but it is uncertain to what extent epilepsy and background EEG oscillations share neurobiological underpinnings. Here, we aimed to assess the genetic correlation between epilepsy and background EEG oscillations.

**Methods**

Confounding factors, including the heterogeneous etiology of epilepsies and medication effects, hamper studies on background brain activity in people with epilepsy. To overcome this limitation, we compared genetic data from a genome-wide association study (GWAS) on epilepsy (n = 12 803 people with epilepsy and 24 218 controls) with that from a GWAS on background EEG (n = 8425 subjects without epilepsy), in which background EEG oscillation power was quantified in four different frequency bands: alpha, beta, delta, and theta. We replicated our findings in an independent epilepsy replication dataset (n = 4851 people with epilepsy and 20 428 controls). To assess the genetic overlap between these phenotypes, we performed genetic correlation analyses using linkage disequilibrium score regression, polygenic risk scores, and Mendelian randomization analyses.

**Results**

Our analyses show strong genetic correlations of genetic generalized epilepsy (GGE) with background EEG oscillations, primarily in the beta frequency band. Furthermore, we show that subjects with higher beta and theta polygenic risk scores have a significantly higher risk of having generalized epilepsy. Mendelian randomization analyses suggest a causal effect of GGE genetic liability on beta oscillations.

**Significance**

Our results point to shared biological mechanisms underlying background EEG oscillations and the susceptibility for GGE, opening avenues to investigate the clinical utility of background EEG oscillations in the diagnostic workup of epilepsy.

# Introduction

The power of oscillations in background electroencephalogram (EEG) is a highly stable and heritable human trait.[1] It is easily acquired and can be automatically analyzed by software, rather than subjective interpretation. Epilepsy is highly heritable and is characterized by altered brain excitability.[2, 3] Oscillatory activity is believed to serve an essential role in corticothalamic functioning, and can be measured as power of oscillations in background EEG at different broadband frequencies.[4] Neurophysiological relationships between background EEG and generalized epileptiform discharges have been well described.[5-8] However, it is currently unknown whether background oscillatory activity is itself associated with epilepsy, and whether background EEG and epilepsy have a shared neurobiological and genetic basis.



There have been some studies where background EEG oscillation measurements have been directly compared between people with epilepsy and healthy controls. However, such studies have yielded conflicting results, most likely because sample sizes were small and antiseizure drugs can strongly affect EEG measurements.[9-14]These limitations and bias can be overcome by large-scale genetic studies, in which genetic determinants of background EEG measurements are assessed independently in healthy controls (presumably not taking antiseizure drugs). These genetic determinants can then be compared to genetic determinants of different epilepsy phenotypes, as assessed in a different study. Comparing these independent studies allows for a well-powered and unbiased assessment of shared genetic determinants of epilepsy and EEG oscillations.

Here, we therefore assessed whether oscillatory background EEG is genetically correlated with focal and generalized epilepsy. The association between genetic variants and background brain activity was previously investigated in a genome-wide association study (GWAS) on 8425 subjects without epilepsy.[15] We combined these data with our recently published large GWAS of epilepsy,[16] to examine genetic correlations between several types of epilepsy and oscillatory brain activity across frequency bands (delta, 1–3.75 Hz; theta, 4–7.75 Hz; alpha, 8–12.75 Hz; and beta, 13–30 Hz). Next, we utilized polygenic risk scoring (PRS) to assess whether people with GGE have a genetic predisposition toward altered background brain activity. We then replicated genetic correlation and polygenic

analyses using an independent cohort from the Epi25 Collaborative (*n* = 4851 people with epilepsy and 20 428 controls). Finally, we performed Mendelian randomization (MR) to gain insight into possible causal relationships between genetic variants associated with epilepsy and those associated with background EEG. We thus provide converging evidence for consistent cross-trait genetic overlap between epilepsy and background EEG.

# Materials and methods

## Study population: Discovery dataset

The participants derived from the epilepsy GWAS[16] for the current analyses were Caucasian subjects. The epilepsy GWAS included 13 control cohorts.[16] Case/control ascertainment and diagnostic criteria were previously reported.[16] As described previously,[16] epilepsy specialists diagnosed people with epilepsy and ascertained phenotypic subtypes. Population-based datasets, some of which had been screened to exclude neurological disorders, were used as controls. However, due to the relatively low prevalence of epilepsy in the general population (~0.5–1%), screening to exclude epilepsy in control cohorts will have only a minor effect on statistical power. Summary statistics from the recent epilepsy GWAS conducted by the International League Against Epilepsy (ILAE) Consortium on Complex Epilepsies GWAS were available for *n* = 12 803 cases (with either focal or generalized epilepsy) and 24 218 controls.[16] From those participants, the following subjects were excluded for those analyses requiring individual-level genotype data: Finnish ancestry (none had genetic generalized epilepsy [GGE]) and the subset of the EPICURE-SP1 cohort that lacked informed consent for the current analyses, resulting in subject-level genotype data being available for 11 446 people with epilepsy and 22 078 controls. Subjects with epilepsy were stratified into GGE (*n* = 3122) and focal epilepsy (*n* = 8324); GGE was further subdivided into childhood absence epilepsy (CAE; *n* = 561), juvenile absence epilepsy (JAE; *n* = 311), juvenile myoclonic epilepsy (JME; *n* = 1000), and generalized tonic–clonic seizures only (GTCS only; *n* = 195). GGE subtype information was not available for 1055 people with epilepsy.

We downloaded summary statistics of the ENIGMA-EEG GWAS of resting state oscillation power in the delta (1–3.75 Hz), theta (4–7.75 Hz), alpha

(8–12.75 Hz), and beta (13–30 Hz) bands at the vertex (Cz) electrode (*n* = 8425 participants).[15] This EEG GWAS was based on five cohorts from four cooperating centers. Although the selection criteria varied across cohorts, all adult cohorts included epilepsy and prolonged unconsciousness after head trauma as exclusion criteria, which were communicated at the time of recruitment or at the first laboratory visit; because neurological disorders were an exclusion criterium, we do not expect subjects to be taking antiseizure drugs (although this was no explicit exclusion criterion). All these were self- or parent-reported retrospective questions. A full sample description and recording specifics are available in the supplement of the original study,[15] and the EEG analysis protocol is available online at http://enigma.ini.usc.edu/ongoing/enigma-eeg-working-group/. In brief, eyes-closed resting EEG was recorded or offline rereferenced to averaged earlobes, visually cleaned with standard criteria by local expert EEG analysts with rogue channels removed, and scanned for sleep transition (eye rolling, alpha dropout). Eye movement was removed using regression or independent component analysis. A minimum of 1 min of recording was required.



Approval for the source studies was obtained by all relevant institutional review boards, and all study participants provided written informed consent according to the Declaration of Helsinki.

## Replication dataset

To replicate our findings, we used data from the Epi25 Collaborative (http://epi-25.org/). This cohort currently comprises 4851 people with epilepsy, of whom 2612 have focal epilepsy and 2239 have GGE (no data on GGE subtypes were available). The cases were matched to a total of 20 428 controls from the Partners Healthcare Biobank (*n* = 14 857), the Epi25 Collaborative (*n* = 210), the Genetics and Personality consortium (*n* = 456), and an in-house project on inflammatory bowel disease (*n* = 4905). The cohorts were genotyped on the Illumina Global Screening Array, with the exception of the Partners Healthcare Biobank participants, who were genotyped on the Illumina Multi-Ethnic Screening Array. Approval was obtained by all relevant institutional review boards, and all study participants provided written informed consent according to the Declaration of Helsinki.

## Genetic correlation analyses

Genetic correlations between epilepsy subtypes and oscillatory brain activity were computed using bivariate linkage disequilibrium score regression (LDSC).[17] For these analyses, as no individual-level genotype data were available from the EEG dataset, we used published summary statistics of the EEG frequency bands (alpha, beta, delta, and theta; $n$ = 8425 participants) and the epilepsy subtypes (focal, GGE, CAE, JAE, JME, and GTCS only; $n$ = 12 803 cases suffering from either focal or generalized epilepsy and 24 218 controls) from the ILAE consortium as a discovery dataset.[16] For LDSC replication analyses, we used unpublished data from the Epi25 Collaborative (http://epi-25.org/; $n$ = 4851 people with epilepsy and 20 428 controls). For discovery and replication LDSC analyses, default settings of LDSC were used, with precomputed linkage disequilibrium (LD) score weights derived from the European subset of the 1000 Genomes project.[18] See Table S1 for the number of single nucleotide polymorphisms (SNPs) per LDSC analysis. The significance threshold was Bonferroni-corrected for the two main epilepsy subtypes studied (GGE and focal) but not for the EEG power spectra, because these were all highly correlated at $p < 10^{-17}$ (Table S2), resulting in a significance threshold of $p$ = .05/2 = .025. Similarly, we did not correct for the individual GGE subtypes, which are phenotypically similar and genetically highly correlated.[16]

## PRS analyses

For PRS analyses, we used individual-level genotype data derived from the epilepsy GWAS[16] and summary statistics from the EEG GWAS.[15] Quality control was performed as reported in the published epilepsy GWAS.[16] We then added a genotype filter for call rate greater than .99 and the exclusion of genetically related subjects to allow for highly conservative PRS estimates. Genetic interrelatedness was calculated with KING,[19] and one subject from each pair with third-degree or higher relatedness (kinship coefficient > .0442) was excluded. PRSice[20] was used with default settings to assess whether subjects with epilepsy had different EEG frequency power PRSs compared to controls. In brief, to each SNP we assigned a weight proportional to its association in the four EEG GWASs (alpha, beta, delta, and theta). Next, individual PRSs were calculated as the sum of weighted effect alleles for every subject from the epilepsy cohort. These PRSs were standardized with a $Z$-score transformation:

$\frac{PRS-mean(PRS)}{SD(PRS)}$. SNPs were pruned to a subset of genetically uncorrelated SNPs (LD $R^2 < .1$), and PRS values were calculated using a number of different *p*-value thresholds from .0001 to .5. Next, logistic regression analyses, corrected for sex and 10 genetic ancestry principal components (PCs), were performed to assess the association of these PRS scores with GGE. The PRS with the highest association with GGE was chosen as the "best fit," after which logistic regression analyses were repeated to assess the association of this PRS with the other epilepsy subtypes. We used a conservative $p < .001$ significance threshold to correct for multiple comparisons, as recommended for PRSice.[20] Explained variance represented by the Nagelkerke $R^2$ was computed using a logistic regression of the PRS, subtracted from the baseline model (covariates only: sex and four PCs). To quantify the association of beta power PRS with GGE, we used PRSice standard settings to divide subjects into 10 deciles based on their beta power PRS scores. We then performed logistic regression to compare the risk of having GGE between every decile, with the lowest (0%–10%) as a reference (corrected for sex and four PCs). We then repeated the analyses in the independent Epi25 cohort. This dataset contained approximately one third fewer GGE cases than the discovery cohort, providing insufficient power to exactly replicate our discovery PRS findings. We therefore performed quasireplication using a one-sample test of the proportion to assess concordance effect directions between discovery and replication PRS analyses, computing *Z*-scores that were converted into *p*-values.

## Mendelian randomization

Two major limitations of observational studies and other types of studies are unmeasured confounding and uncertainties about cause and effect. MR has the potential to overcome these limitations, as MR leverages genetic instruments (most often SNPs) as exposures as well as outcomes. Because SNPs are not influenced by state-dependent factors, MR has the potential to shed light on potential causal mechanisms between two traits; SNPs strongly associated with two or more traits index these traits without confounding. MR can be done in two directions for two given traits, with each MR analysis testing whether one trait has a potential effect on the other. However, here, we could only conduct one-way MR due to lack of genome-wide significant loci in the EEG GWAS. Several MR techniques are

available, and the consensus is that results from different approaches show robustness and consistency of results across methods.

To explore possible causal effects of GGE genome-wide loci (exposure) on EEG background oscillations (outcome), we thus conducted MR analyses using GGE and EEG summary statistics data. Two hundred twenty-eight SNPs significantly associated with GGE ($p < 5 \times 10^{-8}$) were extracted from both the GGE and EEG GWASs. The summary statistics of 228 SNPs were harmonized to ensure the SNP effect direction corresponded with equal effect alleles across GGE and EEG. We used the "TwosampleMR" package[21] in R to perform fixed effects inverse variance-weighted (IVW), weighted median, and MR Egger models. We then performed sensitivity analyses, including horizontal pleiotropic effects estimated by the intercept of MR Egger, residual heterogeneity due to pleiotropy estimated by Cochran Q test,[22] and leave-one-out analyses (for the fixed effects IVW model), to evaluate whether any single instrumental variable was driving the results. Generalized summary data-based MR (GSMR) analyses were performed using the "GSMR"[23] package in R. To that end, first the LD matrix of the selected SNPs was calculated using PLINK[24] and GCTA[25] within 1000 Genomes Phase 3 data.[18] The minimum number of instrumental variables in the GSMR model was loosened from 10 to five as there were only eight independent ($r^2 < .01$, LD window = 10 Mb) significant loci identified in the GWAS of GGE (and none in the EEG GWAS). We used default options in GSMR with heterogeneity in dependent instruments (HEIDI) testing for instrumental outliers' detection. At the end, we repeated GSMR with loosened LD prune thresholds (i.e., $r^2 < .1$, $r^2 < .15$, and $r^2 < .2$), because GSMR takes LD structure into account by adding the LD matrix. The significance threshold was Bonferroni corrected for all seven of these MR models ($p = .05/7 = .007$).

# Results

## Genetic correlations between epilepsy and oscillatory brain activity

In a total study population of 45 446 subjects ($n = 8425$ from the EEG and $n = 37\,021$ from the epilepsy GWASs), we computed genetic correlations (Rg) of alpha, beta, delta, and theta oscillatory brain activity with focal epilepsy and GGE. We

found significant correlations between GGE and beta power (Rg = 0.44 ± SE of .18, $p$ = .01) and theta power (Rg = 0.25 ± 0.11, $p$ = .02; Figure 1, upper panel, Table S1). This was further supported by the correlations between beta power and theta power with the GGE subtypes CAE, JAE, and JME; all had similarly high correlation coefficients. We found no genetic correlations between focal epilepsy and any of the EEG phenotypes. We then attempted to replicate the genetic correlations using the unpublished Epi25 dataset and found genetic correlations similar (in both sign and effect size) to the discovery analyses (Figure 1, lower panel); GGE correlated with beta power (Rg = 0.52 ± 0.21, $p$ = .01), whereas the genetic correlation between theta power and GGE paralleled the discovery cohort (albeit not reaching significance: Rg = 0.16 ± 0.12, $p$ = .18). All genetic correlation estimates with focal epilepsy were again nonsignificant. There were no data available for GGE subtypes.



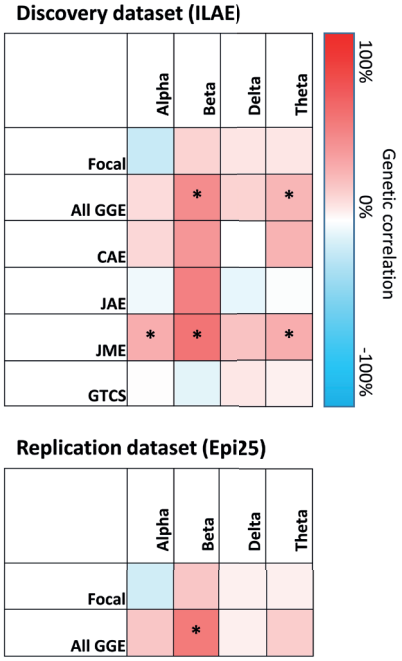



**Figure 1: Genetic correlations between electroencephalographic (EEG) frequency bands and epilepsy subtypes.** Genetic correlations were calculated by comparing the EEG frequency band genome-wide association study (GWAS) with the International League Against Epilepsy (ILAE) GWAS (upper panel, discovery dataset) and the Epi25 GWAS (lower panel, replication dataset). * p < .05. CAE, childhood absence epilepsy; GGE, genetic generalized epilepsy; GTCS, generalized tonic–clonic seizures; JAE, juvenile absence epilepsy; JME, juvenile myoclonic epilepsy.

## Oscillatory brain activity polygenic scores are associated with generalized epilepsy

We used polygenic scoring to utilize the full distribution of background EEG-associated SNPs to assess whether people with epilepsy have a different polygenic score for specific frequency bands compared to controls. We observed significant positive associations between beta and theta power PRSs with GGE, in line with the LDSC results (Figure 2). In particular, beta power PRSs were strongly associated with GGE (beta = .11, SE = .020, $p$ = 5.3 × $10^{-8}$, explained variance = .21%; Figure 2),

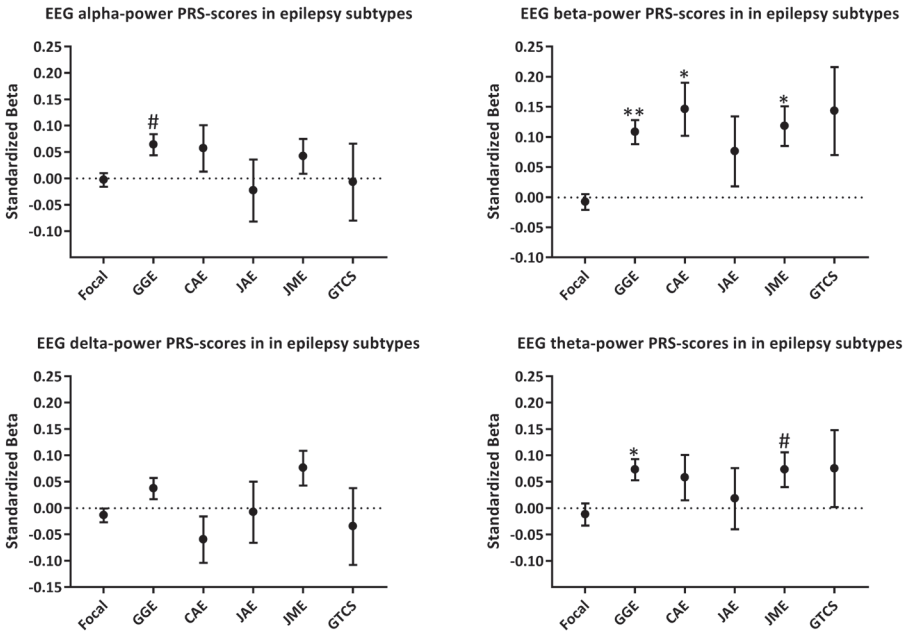



**Figure 2: Beta and theta power electroencephalographic (EEG) oscillation polygenic risk scores (PRSs) are associated with generalized epilepsy but not with focal epilepsy.** The "best-fit" p-value threshold (pt) was chosen based on the most significant association with genetic generalized epilepsy (GGE), which was then applied to all other epilepsy subtypes. The numbers of single nucleotide polymorphisms included in each model were 2670 for alpha power (pt = .0105), 10 861 for beta power (pt = .06245), 8182 for delta power (pt = .0446), and 3833 for theta power (pt = .01665). Logistic regression analyses were performed to assess the association between the PRSs and the different epilepsy subtypes, corrected for sex and 10 principal components. #p < .05, *p < .001, **p < 10-7. Childhood absence epilepsy (CAE), generalized tonic–clonic seizures only (GTCS), juvenile absence epilepsy (JAE), and juvenile myoclonic epilepsy (JME) are GGE subtypes. Focal, focal epilepsy.

which was further supported by significant associations of beta power PRS with its subtypes CAE (beta = .15, SE = .044, $p = 8.5 \times 10^{-4}$) and JME (beta = .12, SE = .033, $p = 3.6 \times 10^{-4}$). Furthermore, of the participants in the GGE case–control cohort, those in the highest 10% decile of beta power PRS scores were 1.4-fold more likely to have GGE compared to the people in the lowest 10% PRS decile (Figure 3; odds ratio [OR] = 1.40, 95% confidence interval [CI] = 1.18–1.67, $p = 1.5 \times 10^{-4}$). When using the independent Epi25 cohort as a replication dataset, we found that the directions of effect agreed with the discovery analyses for all associations between EEG PRSs and GGE ($p_{one-sided} = .023$, $p_{two-sided} = .046$; Figure S1). EEG PRSs were not significantly different between people with focal epilepsy and controls.



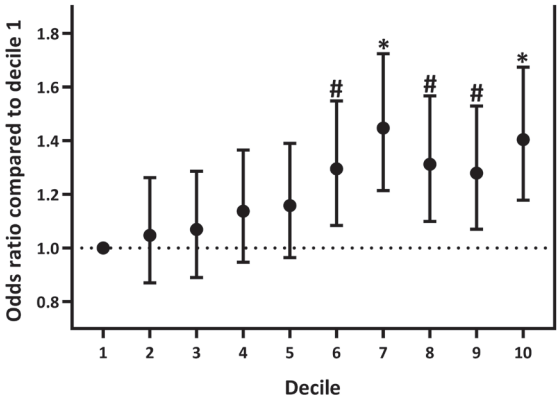



**Figure 3: Polygenic risk score (PRS) analyses show that higher beta-power PRS is associated with an increased likelihood of having GGE.** All subjects were divided into 10 deciles based on their beta-power PRS scores. Logistic regression analyses were performed to quantify the increased risk of having GGE between every decile compared to the lowest decile (0–10%) as a reference. The odds ratios of these analyses are displayed on the Y-axis. #p<.05; *p<.001.

## MR analyses

MR analyses were performed to assess potential causative relationships between background EEG and GGE. Eight GGE-associated SNPs were selected as instrumental variables at a strict LD prune threshold ($r^2 < .01$, LD window = 10 Mb). These were used in fixed effects IVW, weighted median, MR Egger, and GSMR ($r^2 < .01$) models. After loosening the LD threshold, 11 ($r^2 < .1$), 12 ($r^2 < .15$), and 14 ($r^2 < .2$) SNPs were selected as instrumental

variables for GSMR models. Causal effects of GGE loci on beta oscillations were found at the LD $r^2 < .15$ and $r^2 < .2$ thresholds (OR = 1.79, 95% CI = 1.189–2.707, $p = 5.2 \times 10^{-3}$ and OR = 1.723, 95% CI = 1.180–2.516, $p = 4.8 \times 10^{-3}$, respectively; Table S4, Figure S2). Significant heterogeneity was detected in the fixed effects IVW model (Q-statistic = 18.188, $df = 7$, $p = .01$) and MR Egger model (Q-statistic = 14.594, $df = 6$, $p = .02$). No SNPs altered the pooled β coefficient in the leave-one-out sensitivity analysis ($\beta = .374$, $p = .314$) in the fixed effects IVW model. We found no evidence of horizontal pleiotropic effects. Similarly, the HEIDI test detected no SNPs as pleiotropic outliers.

# Discussion

Here, we leveraged the largest currently available GWASs to assess shared genetic underpinnings of epilepsy and of background EEG oscillations. In particular, we found strong genetic relationships between GGE and beta power oscillations, which were replicated in an independent sample.

Previous studies comparing EEG background oscillations between people with epilepsy and controls are inconsistent; some show increased power in all frequency bands (alpha, beta, delta, theta), whereas others show only increases in specific frequency bands or even decreases in power.[9–14] This heterogeneity likely reflects multiple variables that are difficult to control for in clinical studies, such as antiepileptic drug (AED) usage, sleep deprivation, influence of (inter-)ictal epileptic brain activity, EEG processing, and electrode placement. We overcame such limitations by determining the genetic underpinnings of EEG frequency bands in people without epilepsy who are AED-naive, and with consistent electrode placement and signal processing. We applied several statistical models to assess this overlap and found that people with generalized, but not focal, epilepsy carry a relative abundance of genetic variation associated with higher beta oscillations. MR analyses pointed to causal effects of genetic liability to GGE on beta power.

We did not find genetic correlations between background EEG and focal epilepsy. Although power was limited for this analysis, this finding is consistent with the low contribution of common genetic variants in focal

epilepsy and the lack of genetic overlap between focal and generalized epilepsy.[16] Focal epilepsy is likely to represent a more heterogenous group of different causes of epilepsy, many of which do not have a primary genetic cause (e.g., symptomatic epilepsy after traumatic brain injury). Moreover, focal epilepsy by definition only affects one part of the brain and is therefore less likely to be associated with germline genetic variation and background EEG oscillations, which most likely affect the whole brain. Although we found associations of common variants with focal epilepsy in our latest GWAS, the overall polygenic burden and SNP-based heritability was modest compared to GGE.[16] This suggests that further studies assessing common genetic variants in focal epilepsy are less likely to yield major advances. Perhaps further studies on smaller, more homogenous focal epilepsy cohorts or studies assessing rare genetic variants could yield more insights into its pathophysiology. In contrast to focal epilepsy, the EEG discharges that characterize generalized epilepsy are dependent on the thalomocortical system.[5, 26] Similarly, background oscillations have been functionally attributed to the thalamocortical system,[27, 28] suggesting that thalamocortical functioning could represent a common neurobiological mechanism reflecting overall brain excitability, which influences both GGE risk and (beta power) background oscillations.

Our results should be interpreted in the light of several limitations. First, we are aware of the possible advantages of using genome complex trait analysis (GCTA) relative to LDSC, but because no subject-level genotype data are available for the EEG GWAS, we restricted our genetic correlation estimates to LDSC, which is based on summary statistics. LDSC has proven to be a reliable method for genetic correlation estimates, and results between LDSC and GCTA have proven consistent. Second, we found that the same genetic variants underlie both GGE and beta power oscillations, but our study does not prove that people with GGE have altered background oscillations, because we did not have EEG measurements of people with epilepsy in this study. Third, only one-way MR analyses were performed due to lack of genome-wide significant loci in the EEG GWAS. Our results suggest that GGE causally influences beta power oscillations. However, we cannot exclude the possibility of bidirectional causality between EEG

and GGE, and thus it could also be possible that beta power has a causal effect on GGE risk. Fourth, we had insufficient data available to carry out subgroup analyses on subjects with nonlesional focal epilepsy.

Altogether, our results point to shared biological mechanisms underlying background EEG oscillations and the susceptibility for generalized seizures. Our findings thus open avenues to investigate the clinical utility of background oscillations in genetic generalized epilepsy. Potentially, prospective studies could confirm whether altered beta oscillatons could be a prodromal state of GGE or whether aberrant beta oscillations constitute a feature of epilepsy. Future studies may also integrate transcranial magnetic stimulation–EEG and/or event-related potentials to examine whether beta and theta powers correlate with altered brain excitability in subjects with high epilepsy liability. We hypothesize that the genetic correlation between GGE and background oscillations will be reflected by measurable differences in background EEG measures between people with and without GGE, which could be used in the diagnostic workup after a first suspected seizure. This information can be used in machine-learning studies by integrating background EEG with other sources of clinical and demographic data, which may one day increase the accuracy of epilepsy diagnosis.

## Acknowledgements

We are grateful to the people with epilepsy and volunteers who participated in this research. We thank the following clinicians and research scientists for their contribution through sample collection (cases and controls), data analysis, and project support: Geka Ackerhans, Muna Alwaidh, R. E. Appleton, Willem Frans Arts, Guiliano Avanzini, Paul Boon, Sarah Borror, Kees Braun, Oebele Brouwer, Hans Carpay, Karen Carter, Peter Cleland, Oliver C. Cockerell, Paul Cooper, Celia Cramp, Emily de los Reyes, Chris French, Catharine Freyer, William Gallentine, Michel Georges, Peter Goulding, Micheline Gravel, Rhian Gwilliam, Lori Hamiwka, Steven J. Howell, Adrian Hughes, Aatif Husain, Monica Islam, Floor Jansen, Mary Karn, Mark Kellett, Ditte B. Kjelgaard, Karl Martin Klein, Donna Kring, Annie W. C. Kung, Mark Lawden, Jo Ellen Lee, Benjamin Legros, Leanne Lehwald, Edouard Louis, Colin H. T. Lui, Zelko Matkovic, Jennifer McKinney, Brendan McLean, Mohamad Mikati, Bethanie

Morgan-Followell, Wim Van Paesschen, Anup Patel, Manuela Pendziwiat, Marcus Reuber, Richard Roberts, Guy Rouleau, Cathy Schumer, B. Sharack, Kevin Shianna, N. C. Sin, Saurabh Sinha, Laurel Slaughter, Sally Steward, Deborah Terry, Chang-Yong Tsao, T. H. Tsoi, Patrick Tugendhaft, Jaime-Dawn Twanow, Jorge Vidaurre, Sarah Weckhuysen, Pedro Weisleder, Kathleen White, Virginia Wong, Raju Yerra, Jacqueline Yinger, and all contributing clinicians from the Department of Clinical and Experimental Epilepsy at the National Hospital for Neurology and Neurosurgery and University College London Institute of Neurology. We would like to thank the Ming Fund for providing funding for R.S. This work was in part supported by a Translational Research Scholars award from the Health Research Board of Ireland (Christopher D. Whelan) and by research grants from Science Foundation Ireland (16/RC/3948 and 13/CDA/2223), and cofunded under the European Regional Development Fund and by FutureNeuro industry partners. Further funding sources include Wellcome Trust (grant 084730); Epilepsy Society, UK, National Institute for Health Research (NIHR; 08-08-SCC); GIHE, National Institutes of Health (NIH) R01-NS-49306-01 (Russell J. Buono); NIH R01-NS-053998 (Daniel H. Lowenstein); GSCFE, NIH R01-NS-064154-01 (Russell J. Buono, Hakon Hakonarson); NIH UL1TR001070, Development Fund from the Children's Hospital of Philadelphia (Hakon Hakonarson); National Health and Medical Research Council program grant 1091593 (Samuel F. Berkovic, Ingrid E. Scheffer, Karen L. Oliver, Katja E. Boysen); Royal Melbourne Hospital Foundation Lottery Grant (Slavé Petrovski); Royal Melbourne Hospital Neuroscience Foundation (Terence J. O'Brien); European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreements 279062 (EpiPGX) and 602102, Department of Health NIHR Biomedical Research Centres funding scheme, European Community (EC; FP6 project EPICURE: LSHM-CT2006-037315); German Research Foundation (DFG; SA434/4-1/4-26-1 (Thomas Sander), WE4896/3-1); EuroEPINOMICS Consortium (European Science Foundation/DFG: SA434/5-1, NU50/8-1, LE1030/11-1, HE5415/3-1 [Thomas Sander, Peter Nürnberg, Holger Lerche, Ingo Helbig], RO 3396/2-1); German Federal Ministry of Education and Research, National Genome Research Network (NGFNplus/EMINet: 01GS08120, and 01GS08123 [Thomas Sander, Holger Lerche]; IntenC, TUR 09/I10 [Thomas Sander]); Netherlands National Epilepsy Fund (grant 04-08); EC (FP7 project EpiPGX 279062); and

Research Grants Council of the Hong Kong Special Administrative Region, China project numbers HKU7623/08 M (Stacey S. Cherny, Patrick Kwan, Larry Baum, Pak C. Sham), HKU7747/ 07 M (Stacey S. Cherny., Pak C. Sham), and CUHK4466/06 M (Patrick Kwan, Larry Baum). Collection of Belgian cases was supported by the Fonds National de la Recherche Scientifique, Fondation Erasme, Université Libre de Bruxelles. GlaxoSmithKline funded the recruitment and data collection for the GenEpA Consortium samples. We acknowledge the support of Nationwide Children's Hospital in Columbus, Ohio, USA. The Wellcome Trust (WT066056) and the NIHR Biomedical Research Centres Scheme (P31753) supported UK contributions. Further support was received through the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (contract N01HD33348). The project was also supported by the popgen 2.0 network through a grant from the German Ministry for Education and Research (01EY1103). Parts of the analysis of this work were performed on resources of the High Performance Center of the University of Luxembourg and Elixir–Luxembourg. The KORA study was initiated and financed by the Helmholtz Zentrum München–German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences, Ludwig Maximilian University, as part of LMUinnovativ. The ILAE facilitated the Consortium on Complex Epilepsies through the Commission on Genetics and by financial support; however, the opinions expressed in the article do not necessarily represent the policy or position of the ILAE.

## Conflict of interest

None of the authors has any conflict of interest to disclose.

## Author contributions

R.S., J.J.L., D.Sm., and B.P.C.K. contributed to the conception and design of the study. R.S., J.J.L., B.D.L., D.Sm., and B.P.C.K. contributed to the acquisition and analysis of data. R.S., J.J.L., B.D.L., C.L., D.L., A.S., D.Sc., J.A.C., K.R., R.A.R.B., O.D., K.P.J.B., F.E.J., D.Sm., and B.P.C..K. contributed

to the drafting of the manuscript and preparing the figures. Members of the ILAE Consortium on Complex Epilepsies and EPI25 Collaborative contributed clinical and genetic data.

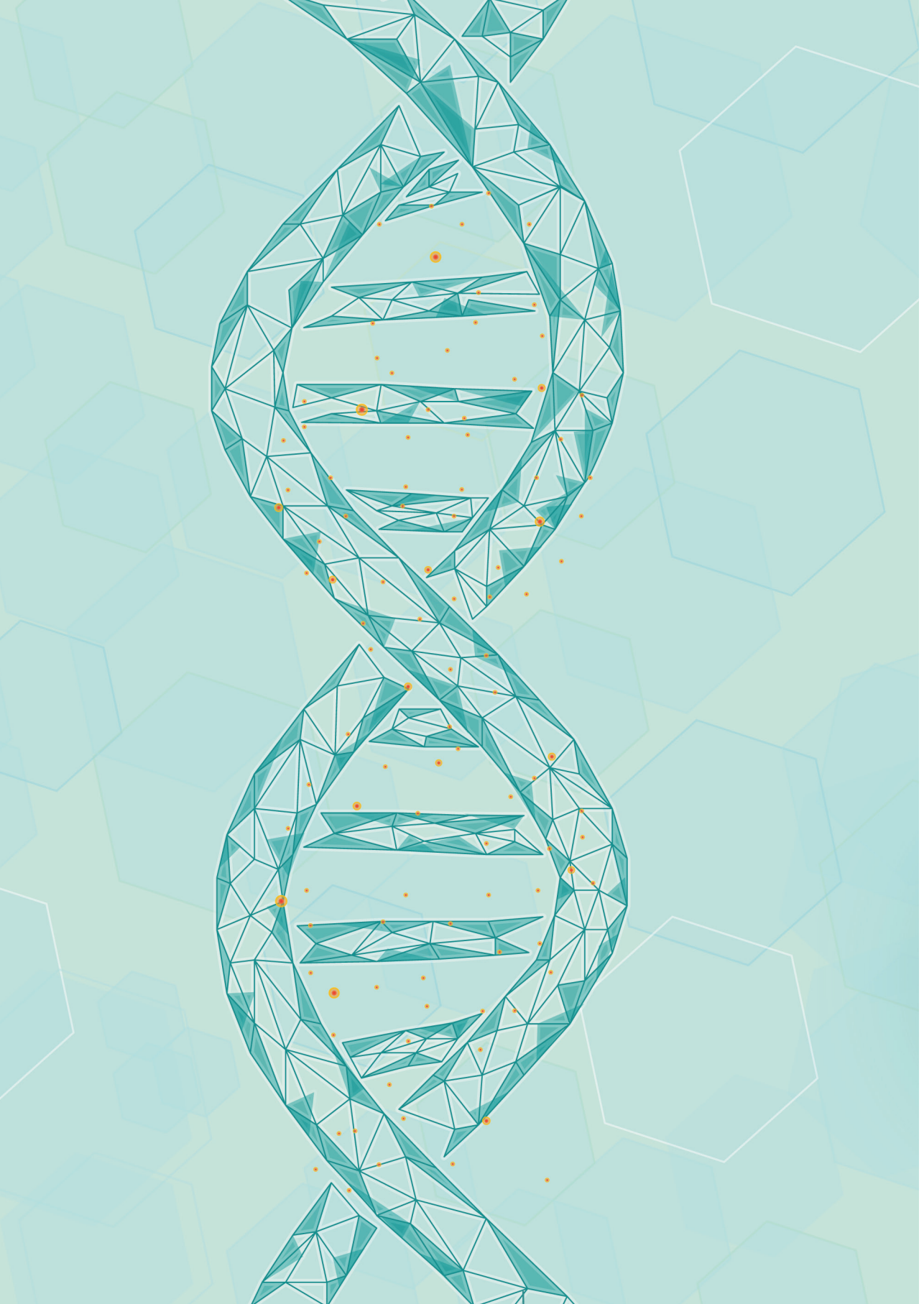
## Supporting information

Additional supporting information may be found at: https://tinyurl.com/4j3srm5j.

# References


1. Smit DJA, Posthuma D, Boomsma DI, De Geus EJC. Heritability of background EEG across the power spectrum. Psychophysiology. 2005;42:691– 7.
2. Devinsky O, Vezzani A, O'Brien TJ, Jette N, Scheffer IE, de Curtis M, et al. Epilepsy. Nat Rev Dis Primers. 2018;4:18024.
3. Thomas RH, Berkovic SF. The hidden genetics of epilepsy— a clin-ically important new paradigm. Nat Rev Neurol. 2014;10:283– 92.
4. Steriade M. Corticothalamic resonance, states of vigilance and mentation. Neuroscience. 2000;101:243– 76.
5. Gloor P. Generalized epilepsy with bilateral synchronous spike and wave discharge. New findings concerning its physiologi-cal mechanisms. Electroencephalogr Clin Neurophysiol Suppl. 1978;34:245– 9.
6. Gloor P. Generalized epilepsy with spike- and- wave discharge: a reinterpretation of its electrographic and clinical manifestations. The 1977 William G. Lennox Lecture, American Epilepsy Society. Epilepsia. 1979;20:571– 88.
7. Gloor P, Fariello RG. Generalized epilepsy: some of its cellular mechanisms differ from those of focal epilepsy. Trends Neurosci. 1988;11:63– 8.
8. Steriade M. Sleep, epilepsy and thalamic reticular inhibitory neu-rons. Trends Neurosci. 2005;28:317– 24.
9. Clemens B, Szigeti G, Barta Z. EEG frequency profiles of idiopathic generalised epilepsy syndromes. Epilepsy Res. 2000;42:105– 15.
10. Clemens B. Pathological theta oscillations in idiopathic gener-alised epilepsy. Clin Neurophysiol. 2004;115:1436– 41.
11. Clemens B, Bessenyei M, Piros P, Tóth M, Seress L, Kondákor I. Characteristic distribution of interictal brain electrical activity in idiopathic generalized epilepsy. Epilepsia. 2007;48:941– 9.
12. Clemens B. Valproate decreases EEG synchronization in a use- dependent manner in idiopathic generalized epilepsy. Seizure. 2008;17:224– 33.
13. Clemens B, Puskás S, Besenyei M, Kovács NZS, Spisák T, Kis SA, et al. Valproate treatment normalizes EEG functional connectiv-ity in successfully treated idiopathic generalized epilepsy patients. Epilepsy Res. 2014;108:1896– 903.
14. Tikka SK, Goyal N, Umesh S, Nizamie SH. Juvenile my-oclonic epilepsy: clinical characteristics, standard and quantitative electroencephalography analyses. J Pediatr Neurosci. 2013;8:97– 103.
15. Smit DJA, Wright MJ, Meyers JL, Martin NG, Ho YYW, Malone SM, et al. Genome-wide association analysis links multiple psy-chiatric liability genes to oscillatory brain activity. Hum Brain Mapp. 2018;39:4183– 95.
16. International League Against Epilepsy Consortium on Complex Epilepsies. Genome- wide mega- analysis identifies 16 loci and s diverse biological mechanisms in the common epilepsies. Nat Commun. 2018;9(1):5269.
17. Bulik- Sullivan BK, Loh P- R, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD score regression distinguishes confounding from polygen-icity in genome-wide association studies. Nat Genet. 2015;47:291– 5.

18. 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.
19. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome- wide association studies. Bioinformatics. 2010;26:2867– 73.
20. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. Bioinformatics. 2015;31:1466– 8.
21. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7:e34408.
22. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol. 2015;30:543– 52.
23. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors and common diseases in-ferred from GWAS summary data. Nat Commun. 2018;9:224.
24. Purcell S, Neale B, Todd- Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole- genome associ-ation and population- based linkage analyses. Am J Hum Genet. 2007;81:559– 75.
25. 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.
26. Steriade M, Contreras D. Spike- wave complexes and fast com-ponents of cortically generated seizures. I. Role of neocortex and thalamus. J Neurophysiol. 1998;80:1439– 55.
27. Steriade M, Amzica F, Contreras D. Synchronization of fast (30– 40 Hz) spontaneous cortical rhythms during brain activation. J Neurosci. 1996;16:392–417.
28. Steriade M, Contreras D, Amzica F, Timofeev I. Synchronization of fast (30– 40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks. J Neurosci. 1996;16:2788– 808.



## CHAPTER 5

# USING COMMON GENETIC VARIANTS TO FIND DRUGS FOR COMMON EPILEPSIES

Nasir Mirza, Remi Stevelink, Basel Taweel, Bobby P C Koeleman, and Anthony G Marson, International League Against Epilepsy Consortium on Complex Epilepsies

*Brain Communications.* 2021; Dec 4;3(4):fcab287.

## Abstract

Better drugs are needed for common epilepsies. Drug repurposing offers the potential of significant savings in the time and cost of developing new treatments. In order to select the best candidate drug(s) to repurpose for a disease, it is desirable to predict the relative clinical efficacy that drugs will have against the disease. Common epilepsy can be divided into different types and syndromes. Different antiseizure medications are most effective for different types and syndromes of common epilepsy. For predictions of antiepileptic efficacy to be clinically translatable, it is essential that the predictions are specific to each form of common epilepsy, and reflect the patterns of drug efficacy observed in clinical studies and practice. These requirements are not fulfilled by previously published drug predictions for epilepsy. We developed a novel method for predicting the relative efficacy of drugs against any common epilepsy, by using its Genome-Wide Association Study summary statistics and drugs' activity data. The methodological advancement in our technique is that the drug predictions for a disease are based upon drugs' effects on the function and abundance of proteins, and the magnitude and direction of those effects, relative to the importance, degree and direction of the proteins' dysregulation in the disease. We used this method to predict the relative efficacy of all drugs, licensed for any condition, against each of the major types and syndromes of common epilepsy. Our predictions are concordant with findings from real-world experience and randomized clinical trials. Our method predicts the efficacy of existing antiseizure medications against common epilepsies; in this prediction, our method outperforms the best alternative existing method: area under receiver operating characteristic curve (mean ± standard deviation) $0.83 \pm 0.03$ and $0.63 \pm 0.04$, respectively. Importantly, our method predicts which antiseizure medications are amongst the more efficacious in clinical practice, and which antiseizure medications are amongst the less efficacious in clinical practice, for each of the main syndromes of common epilepsy, and it predicts the distinct order of efficacy of individual antiseizure medications in clinical trials of different common epilepsies. We identify promising candidate drugs for each of the major syndromes of common epilepsy. We screen five promising predicted drugs in an animal model: each exerts a significant dose-dependent effect upon seizures. Our predictions are a novel resource for selecting suitable candidate drugs that could potentially be repurposed for each of the major syndromes of common epilepsy. Our method is potentially generalizable to other complex diseases.

# Introduction

A total of 50 million people are affected by epilepsy.[1] Current drug treatments for epilepsy fail to control seizures in~30% of patients[2,3] and cause adverse effects in ~88% of patients[4,5]; ~20% of people with newly diagnosed epilepsy discontinue their first antiseizure medication (ASM) because of intolerable adverse effects.[6] Hence, there is a need for new ASMs with higher efficacy and/or lower toxicity. Drug repurposing—treating a disease using drugs already licensed for other conditions—offers the potential of significant savings in the time and cost of developing new therapies. Numerous drugs licensed for other conditions have the potential of antiepileptic efficacy.[7] In order to select the best candidate drug(s) to repurpose for epilepsy, it is desirable to predict the relative clinical efficacy that drugs will have in people with epilepsy. One established strategy for discovering potentially effective drugs is to, first, identify the proteins that underlie a disease and, then, identify the drugs that affect the disease-proteins. In such analyses, genes associated with a disease are routinely used as proxies for disease-proteins.[8]



Genetic factors can contribute to the development of epilepsies, either as single-gene mutations in rare monogenic epilepsies, or as multiple genetic variants in common epilepsies.[9] Common epilepsies are complex traits with a polygenic origin, which means that the combined effect of many common risk variants contributes to their genetic risk.[9] Common epilepsies are divided into different types and syndromes[10]; for brevity, we use 'forms' as a general term for both types and syndromes. Different forms of common epilepsy have important differences in their genetic determinants,[11] clinical manifestations and response to medications.[12] Hence, to be most useful for common epilepsies, methods of drug prediction must use the specific genes/proteins underlying a particular form of common epilepsy, to make drug predictions that are specific for that particular form of common epilepsy. This has not been achieved by any of the published drug prediction studies for epilepsy.[11,13–17] Some studies have pooled genes/proteins associated with different forms of epilepsy (including rare epilepsies), to produce a single list of drug predictions for all forms of epilepsy[15–17]; these methods are not readily adaptable to individual common epilepsies, as they require a large number of

genes/proteins definitively associated with a disease. Other studies have used genome-wide transcriptomic analysis of human brain tissue from epilepsy surgery[14,15]; such tissue is only available for a very limited number of epilepsy syndromes, and its analysis is hindered by the lack of suitable control brain tissue that is comparable, normal and has been exposed to ASMs. Of course, any transcriptomic changes detected in epileptic brain tissue could be a consequence, rather than a cause, of disease.

The Genome-Wide Association Study (GWAS) is becoming an increasingly powerful tool for revealing the distinct genetic determinants of different common epilepsies.[11,18–20] GWAS results are routinely used to predict new candidate drugs for complex diseases. In the standard approach, significant variants from the GWAS are mapped to genes; drugs that are known to affect the (protein products) of the genes, are predicted to affect the disease.[21] This simplistic approach has a number of methodological deficiencies. It reflects neither the polygenicity of common diseases, nor the polypharmacology of common drugs. It ignores drugs' effects on disease-protein abundance, even though, in order to exert their therapeutic effect, drugs rectify the activity of disease-proteins by modulating their function or abundance or both.[22–24] It disregards the magnitude and direction of change in disease-proteins' activity, and drugs' effects upon it. Potential causal variants below the genome-wide disease significance threshold are ignored. Practically, it produces an unordered and unranked pool of drug names, with no indication of the relative predicted efficacy of the compounds, to enable selection of the most promising candidates. Ultimately, it is liable to producing poor results. Some limitations of the standard approach are addressed by recently developed enhanced techniques for using GWAS results to identify effective drugs,[25–28] but these newer methods and their drug predictions for common epilepsy still leave room for improvement. None of the existing methods make drug predictions for a disease based upon drugs' effects on the function and abundance of proteins, and the magnitude and direction of those effects, relative to the importance, degree and direction of the proteins' dysregulation in the disease. Our aim was to develop such a method, and to use this method to predict the relative efficacy of drugs for each of the major types and syndromes of common epilepsy, and to make

our predictions available as a novel resource for selecting suitable candidate drugs that could potentially be repurposed for each of the major types and syndromes of common epilepsy.

# Materials and methods

Methods are summarized below; further details can be found online in the Supplementary methods.

## Overview



The common epilepsies are divided into different types, which are further subdivided into different syndromes.[10] In the current work, we included the main types and syndromes analysed in the most recent epilepsy GWAS[11]:

i. All epilepsy, which is comprised of generalized, focal and unclassified epilepsies
ii. The two main types of all epilepsy: generalized epilepsy (GE) and focal epilepsy (FE)
iii. Two GE syndromes: juvenile myoclonic epilepsy (JME) and childhood absence epilepsy (CAE)
iv. A FE syndrome: FE with hippocampal sclerosis (HS)

## Method summary

Genetic variants cause disease by modifying the function or abundance (or both) of proteins derived from the variant genes.29 Drugs exert a therapeutic effect by rectifying the abnormal function or abundance (or both) of the proteins underlying a disease.22–24 To predict the relative efficacy of drugs against a disease, we developed (Fig. 1) a novel score for drugs' relative ability to affect the protein function and abundance changes caused by common genetic variations associated with the disease: the disease–protein function and abundance modulation (FAM) score.

For method development and benchmarking, we used the all epilepsy GWAS. Then, we applied the developed method to the GWAS for specific epilepsy types and syndromes.

It should be noted that, to aid brevity and readability, we use the expressions 'disease-associated proteins' and 'disease-proteins' as proxies for 'proteins encoded by genes bearing variations associated with the disease', and we use the expression 'protein abundance changes' as proxy for 'changes in gene expression'.

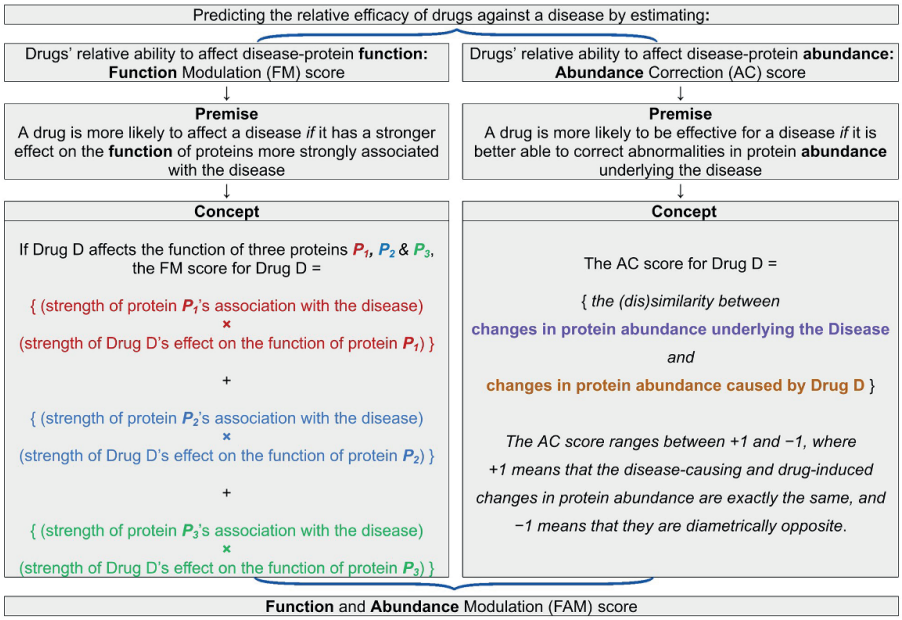



**Figure 1: Premise and conceptual explanation of the disease-protein function modulation (FM) and abundance correction (AC) scores, which are integrated to form the disease-protein function and abundance modulation (FAM) score.** Before integration, the FM score is adjusted to control for the different number of proteins affected by each drug (see Supplementary material for details). Cosine distance is the (dis)similarity metric used for calculating the AC score.

## The disease-protein FAM score: creation and benchmarking

The steps taken in developing the method for calculating the FAM score are detailed in Supplementary material. Below, we summarize the method (Fig. 1) we developed for calculating the FAM score.

The FAM score is calculated by aggregating its two constituent scores:

i. The disease-protein function modulation (FM) score
ii. The disease-protein abundance correction (AC) score

### *FM score*

The FM score is based on the following premise: A drug's ability to affect a disease can be predicted from:

i. the degree of disease-association of each protein whose function is affected by the drug, and
ii. the strength of the drug's effect on the function of each of those proteins

The degree of disease-association of proteins is derived from GWAS gene-based P-values. The strength with which drugs affect proteins' function is derived from drug-target affinity data. Figure 1 presents a conceptual explanation of how the FM score is calculated from these two types of data. A more detailed explanation can be found in the Supplementary material.



### *AC score*

The AC score is based upon the following premise: A drug is more likely to be effective for a disease if it is better able to rectify the protein abundance changes underlying the disease.[30] Disease- and drug-induced transcriptomes were compared in order to predict each drug's relative ability to rectify disease-associated protein abundance changes, as previously described[11] and detailed in the Supplementary material. Briefly, the AC score for a drug is calculated as follows: For each disease-associated protein, the algorithm compares the magnitude and direction of change in the protein's abundance found in the disease, with the magnitude and direction of change in the protein's abundance caused by the drug. Then, drugs are ranked in accordance with their overall predicted corrective effect on the abundance of all disease-associated proteins. To measure the overall effect, a metric called 'cosine distance' is used.[31]

### *Aggregating the FM and AC sores to generate the FAM score*

The FM and AC scores were converted into their respective z-scores. The FAM score for each is calculated by averaging its FM and AC z-scores (see Supplementary material for details).

## Comparing our method with existing alternative advanced methods

We compared our results with the results from two existing and contrasting advanced methods for GWAS-based drug predictions.

### *Network-based method*

An approach employed in a number of studies is to identify the drugs that target genome-wide significant disease-proteins and, in addition, the drugs that target the proteins interacting with genome-wide significant disease-proteins.[32–34] We used the GUILDify v2.0 Web Server[35] to identify such drugs.

### *Gene-set analysis method*

In this method,[36] GWAS gene-based P-values are first converted to z-statistics and, then, a single-sided two-sample t-test is used to determine if the mean z-statistic of the genes that are altered in function by a drug is lower than the mean z-statistic of the genes that are not.

## Validation of the FAM score

For *in silico* validation of the FAM score, we examined the following hypotheses:

- The FAM score for all epilepsy specifically prioritizes the drugs that are effective in people with epilepsy: when drugs are ranked by their FAM score for all epilepsy, drugs used to treat epilepsy are ranked higher than drugs used to treat any other human disease
- The FAM score predicts which ASMs are more clinically effective, and which ASMs are less clinically effective, for each common epilepsy syndrome studied
- The FAM score predicts the observed patterns of relative efficacy of individual clinically-effective ASMs for each common epilepsy syndrome studied

The above hypotheses are further detailed in Results and in Supplementary methods.

To test the above hypotheses, we used the following metrics:

1. Identification of effective drugs: we used area under receiver operated characteristics curve (AUROC) analysis to determine the accuracy with drugs' scores discriminate ASMs from all other drugs, or discriminate more from less clinically-effective subsets of ASMs. AUROC was calculated using the package PRROC (version 1.3.1)[37] in R (version 3.4.3). In assessing the discrimination of ASMs from all other drugs, there is a marked class imbalance, because a very small fraction of all drugs are ASMs. To correct for this imbalance, we employed the standard technique of random under-sampling, which is commonly used in published studies (see Supplementary material for further details and references). Specifically, AUROC was calculated using the set of ASMs and a randomly selected set of other drugs equal in number to the ASMs. This process was repeated 1000 times, and mean (± standard deviation) AUROC was calculated. When discriminating more from less effective ASMs, class imbalance is not an issue and, hence, random under-sampling was not employed.
2. Prioritization of effective drugs: amongst all the drug predictions for a phenotype, we determined the average rank of ASMs, or compared the average rank of more clinically-effective and less clinically-effective ASMs. To ease conceptualization and interpretation of results, we converted ranks to percentile ranks. For example, a drug with a percentile rank of 90 is ranked higher/better than 90% of all drugs. Like numerous published studies, we used the median in order to compute the average of ranks, as it is less liable to skewing by outliers (see Supplementary material for further details and references).

## Statistical analysis

We determined the statistical significance of drug identification and prioritization results by comparing the results to those from a null distribution generated by performing 106 random permutations of the scores assigned to drugs.

## Determining whether the drug predictions are driven by individual highly disease-associated proteins

For each epilepsy, FAM scores were re-calculated after excluding, one at a time, the top 10 most strongly disease-associated proteins (Supplementary Table 3). Drug ranks obtained after excluding a protein were compared with the original drug ranks, using Kendall's $\tau$. Kendall's $\tau$ is a commonly used measure of rank correlation.[38] Kendall's $\tau$ ranges from +1 to −1, where +1 means that two ranked lists are identical to each other, −1 means that they are the exact inverse of each other, and 0 means that there is no relationship between them.

Further details about this analysis can be found in the Supplementary material.

## Top candidate drugs

To aid the selection of suitable candidate drugs for experimental validation and clinical evaluation, we demarcated the most promising candidate drugs for each phenotype: the topmost drug predictions with the greatest enrichment of (more) effective ASMs for that phenotype. A manually curated selection of top candidate drugs for different forms of common epilepsy was also produced.

## Testing top candidate compounds in an animal model

As we used complex genetic data to make our drug predictions, we used a complex genetic model to test our drug predictions. We used a rodent model with a complex genetic seizure disorder[39–42] that manifests as audiogenic generalized seizures: the DBA/2 mouse. We tested the five most highly ranked predictions for GE, after filtering out known ASMs, compounds with existing published evidence in the DBA/2 mouse model, drugs lacking evidence of blood–brain barrier permeability, drugs lacking evidence of safe long-term oral use in humans, compounds insoluble in water or saline and 'controlled substances' that require exceptional legal authorization for procurement under the laws of France, where the animal experiments were performed by a contract research organization.

The animal experiment protocol followed the method described by Dürmüller et al.[43] The study was conducted in compliance with Animal Health regulations, in particular:

i. Council Directive No. 2010/63/UE of 22 September 2010 on the protection of animals used for scientific purposes and French decree No. 2013-118 of 1 February 2013 on the protection of animals;
ii. Porsolt facility accreditation for experimentation (E 53 1031, renewed on 19 April 2016);
iii. The recommendations of the Association for Assessment and Accreditation of Laboratory Animal Care of which the accreditation was granted in June 2012 and renewed in 2018.



Porsolt has an in-house ethics programme, which covers animal care and use within the facility.

Additional experimental details about the animal model testing can be found online in the Supplementary material.

## Code availability

The R code for computing FM and FAM scores is available at https://figshare.com/projects/Using_common_variants_to_find_drugs_for_common_epilepsies/78330. The code is for non-commercial use only.

## Data availability

The following datasets are available for download from the project's data repository page (https://figshare.com/projects/Using_common_variants_to_find_drugs_for_common_epilepsies/78330):

i. GWAS gene-based and tissue-wide association study (TWAS) datasets used in our analyses.
ii. Ranked list of the top predicted drugs for each phenotype.
iii. Our complete set of predictions, listing each drug and its FAM score, for each phenotype.

# Results

## The standard method is inadequate for predicting drugs effective against common epilepsies

In the standard method, drugs are predicted to be efficacious if they modulate the function of proteins that are associated with the disease, according to the GWAS, at a genome-wide level of significance.[21] For all epilepsy, GE and FE, *SCN1A* is the only gene that both (i) reaches genome-wide level of disease-significance, and (ii) produces a protein that is known to be altered in function by any existing compound. For CAE, JME and HS, there are no genes that both (i) reach genome-wide level of disease-significance and (ii) produce a protein that is known to be altered in function by any existing compound. Predicting candidate compounds for epilepsy based upon their ability to affect the function of sodium channel protein Type 1 subunit alpha (the protein product of *SCN1A*) yields a recall (from amongst all ASMs, the fraction predicted to be effective) of 35% and precision (from amongst all drugs predicted to be effective, the fraction that are ASMs) of 32%, which equates to an F-score (harmonic mean of the precision and recall) of 33%. The standard method of drug prediction produces an unordered and unranked set of candidate drugs, with no metrics for the relative predicted efficacy of the candidate compounds. This precludes method evaluation based upon predicted drug rankings and AUROC, and hampers the selection of the most promising candidate drugs for experimental validation. The same set of ASMs is predicted to be effective for the two divergent phenotypes of GE and FE, even though some seizure types in the former are aggravated by the ASMs that are most effective for the latter. Hence, for different common epilepsies, this method either fails to identify the majority of known effective drugs, or identifies no candidate drugs at all, or identifies potentially aggravating drugs. By extension, applying the standard approach to common epilepsies will yield no or few candidates for repurposing, will not prioritize amongst the candidates, will fail to identify any or most of the efficacious compounds and will potentially identify aggravating drugs.

## Creating and benchmarking a new method for predicting the relative efficacy of drugs against common epilepsies

To predict the relative efficacy of drugs against common epilepsies, we devised the disease-protein FAM score, which is calculated using the method illustrated in Fig. 1.

For benchmarking, we used the FAM score and alternative existing advanced methods to predict drugs for all epilepsy, and compared the methods' performance. For the identification and prioritization of ASMs, the FAM score achieved AUROC (mean ± standard deviation) of 0.83 ± 0.03 and average percentile of 94, respectively. In comparison, the best performing alternative method achieved AUROC (mean ± standard deviation) of 0.63 ± 0.04 and average percentile of 77. Results of all comparator alternative approaches are shown in Supplementary Table 1.



## Validating the FAM score

Next, we present results of the analyses performed to test the validity of the predictions made using the FAM score.

## The FAM score for all epilepsy specifically prioritizes the drugs that are effective in people with epilepsy

When drugs are ranked by their FAM score for all epilepsy, drugs used to treat epilepsy are ranked higher than drugs used to treat any other human disease. The median rank of drugs used to treat epilepsy is at least seven percentiles higher than that of drug-sets used to treat other human diseases. Permutation-based $P$-value = $1 \times 10^{-4}$ that ASMs are ranked highest, and so much higher than all other drug-sets used to treat all other human diseases.

## The FAM score predicts which ASMs are *more* clinically effective, and which ASMs are *less* clinically effective, for each common epilepsy syndrome

Different ASMs are most effective for different syndromes of common epilepsy. Clinical studies and experience show that, for each common

epilepsy syndrome, some ASMs can be classified into a more clinically-effective subset and some into a less clinically-effective subset. For each common epilepsy syndrome, the FAM score predicts which ASMs are amongst the more efficacious in clinical practice, and which ASMs are amongst the less efficacious in clinical practice (Table 1). Specifically, for each common epilepsy syndrome, the FAM score (i) distinguishes the more from the less clinically-effective ASMs and (ii) prioritizes the more clinically-effective ASMs higher than the less clinically-effective ASMs (Table 1).

| Epi | Identification of AEDs (AUROC) | | | Prioritisation of AEDs (average percentile) | | *P* |
|---|---|---|---|---|---|---|
| | More effective AEDs from all drugs (mean ± SD) | Less effective AEDs from all drugs (mean ± SD) | More from less effective AEDs | More effective AEDs | Less effective AEDs | |
| **HS** | 0.65 ± 0.13 | 0.36 ± 0.18 | 0.87 | 73 | 27 | $8 \times 10^{-3}$ |
| **GE** | 0.85 ± 0.04 | 0.69 ± 0.09 | 0.71 | 93 | 70 | $<1 \times 10^{-6}$ |
| **JME** | 0.88 ± 0.04 | 0.76 ± 0.08 | 0.72 | 96 | 86 | $<1 \times 10^{-6}$ |
| **CAE** | 0.75 ± 0.05 | 0.45 ± 0.15 | 0.79 | 85 | 48 | $2.9 \times 10^{-5}$ |

**Table 1: Performance of the FAM score, measured by the identification and prioritization of AEDs.** Constituents of the 'More effective AEDs' and 'Less effective AEDs' drug-sets are specific to each phenotype. 'Less effective AEDs' comprise the set of less effective, ineffective or aggravating AEDs for that phenotype. AUROC is calculated using drugs' FAM scores. AUROC for identifying AEDs from all drugs is computed using the technique of random under-sampling, and presented as mean ± standard deviation (see Supplementary methods). Prioritization is calculated using drugs' ranks, when all drugs have been ranked from highest to lowest predicted effect on the phenotype. Prioritization result shown is the average (median) rank of AEDs, expressed as a percentile; it is equivalent to the percentage of all drugs ranked below the middle-ranked AED (see Supplementary methods). AUROC, area under the receiver operating characteristics; CAE, childhood absence epilepsy; Epi, epilepsy type or syndrome; GE, generalized epilepsy; HS, focal epilepsy with hippocampal sclerosis; JME, juvenile myoclonic epilepsy; P, permutation-based P-value after Benjamini–Hochberg correction; SD, standard deviation.

In order to predict which ASMs are more clinically-effective and which ASMs are less clinically-effective for a syndrome, the best results are obtained by using the FAM score for that syndrome. To illustrate this, we show that the ASMs that are more effective for CAE are favoured over the ASMs that are less effective for CAE, only when drugs are predicted using the FAM scores for CAE (AUROC: 0.79), and not when drugs are predicted using the FAM scores for all epilepsy, GE, JME, FE or HS (max AUROC:

0.49); permutation-based *P*-value = $1 \times 10^{-5}$ that the AUROC values for CAE and other phenotypes are so contrasting.

For FE, current ASMs are not readily classified into more clinically-effective and less clinically-effective subsets. The FE FAM score identifies and prioritizes ASMs: AUROC (mean ± standard deviation) of $0.85 \pm 0.03$ and average percentile of 94; the FAM score' performance is statistically significant (permutation-based *P*-value = $1 \times 10^{-6}$), and superior to that of its constituent scores.

When considering the ability to distinguish more effective ASMs from all drugs and from less effective ASMs, the FAM score outperforms its constituent scores (Supplementary Table 2).



## The FAM score predicts the observed patterns of relative efficacy of individual clinically-effective ASMs

We tested our predictions against the following observed patterns of relative efficacy of individual clinically-effective ASMs.

## Valproate is the most effective ASM for GE, whereas carbamazepine is the most effective ASM for FE

It is recognized that the efficacy of valproate for generalized onset seizures is 'unsurpassed',[44] whist for focal onset seizures, 'no other drug has been shown to be more effective' than carbamazepine.[45] In our predictions for GE, valproate is ranked highest of all current ASMs. In our predictions for FE, carbamazepine is ranked highest of all current ASMs. Valproate and carbamazepine are amongst the top two of all drugs in our predictions for GE and FE, respectively; permutation-based *P*-value = $5.6 \times 10^{-6}$ for both valproate and carbamazepine being ranked so highly in our predictions for GE and FE, respectively.

## The predicted order of efficacy of ASMs for FE matches that seen in the SANAD trials

The SANAD studies are the largest published head-to-head comparison of multiple ASMs for FE, and the largest published randomized controlled trial of ASMs for FE.[46,47]

Five ASMs were compared in the FE arm of SANAD I: carbamazepine, gabapentin, lamotrigine, oxcarbazepine and topiramate. These drugs' predicted order of efficacy for FE matches the observed order of efficacy in the SANAD I trial. The finding that these drugs are ranked as highly and in the correct order is unlikely to occur by chance ($P < 1 \times 10^{-6}$ by permutation).

Carbamazepine and gabapentin are effective ASMs but, in the FE arm of the SANAD I trial, carbamazepine was significantly more efficacious than gabapentin. Carbamazepine and gabapentin are ranked high in our predictions for FE (percentile ranks 100 and 79, respectively), but carbamazepine is ranked significantly higher than gabapentin (permutation-based *P*-value = $1 \times 10^{-4}$ for the ranks of both drugs being as high but as disparate as observed).

The ASMs compared in the FE arm of SANAD II were lamotrigine, levetiracetam and zonisamide. These drugs' predicted order of efficacy for FE matches the observed order of efficacy in the SANAD II trial. The finding that these drugs are ranked as highly and in the correct order is unlikely to occur by chance ($P < 1 \times 10^{-6}$ by permutation).

## The prioritized order of efficacy of ASMs for GE matches that seen in the SANAD I trial

The SANAD studies are the largest published head-to-head comparison of multiple ASMs for GE, and the largest published randomized controlled trial of ASMs for GE.[48,49]

The ASMs compared in the GE arm of SANAD I were lamotrigine, topiramate and valproate. These drugs' predicted order of efficacy matches the clinically observed order of efficacy in the SANAD I trial. The finding that these drugs are ranked as highly and in the correct order is unlikely to occur by chance (permutation-based *P*-value = $1 \times 10^{-5}$).

Valproate and lamotrigine are effective ASMs but, in the GE arm of the SANAD I trial, valproate was significantly more efficacious than lamotrigine. Valproate and lamotrigine are ranked high in our predictions for GE (percentile ranks 100 and 81, respectively), but valproate is ranked

significantly higher than lamotrigine (permutation-based *P*-value = $3 \times 10^{-4}$ for the ranks of both drugs being as high but as disparate as observed).

The ASMs compared in the GE arm of SANAD II were levetiracetam and valproate. Valproate and levetiracetam are effective ASMs but, in the GE arm of the SANAD II trial, valproate was significantly more efficacious than levetiracetam. Valproate and levetiracetam are ranked high in our predictions for GE (ranks 1 and 15, respectively), but valproate is ranked significantly higher than levetiracetam (permutation-based *P*-value $< 1 \times 10^{-5}$ for the ranks of both drugs being as high but as disparate as observed).



## Topiramate is more effective than lamotrigine for GE, but lamotrigine is more effective than topiramate for FE, in concordance with the SANAD I trial

Lamotrigine and topiramate are the only two ASMs included in both the FE and GE arms of the SANAD I study. In the GE arm of SANAD I, topiramate was more efficacious than lamotrigine, whereas in the FE arm, lamotrigine was more efficacious then topiramate. In our predictions for FE, lamotrigine is ranked higher than topiramate, while for GE, topiramate is ranked higher than lamotrigine. The contrasting ranks of lamotrigine and topiramate for FE and GE are unlikely to occur by chance (permutation-based *P*-value = $1 \times 10^{-4}$).

## For JME, valproate is most effective

Valproate is thought to be the most efficacious broad-spectrum ASM for JME[50–52] but this is based on anecdotal data and retrospective analyses. Amongst our predictions for JME, valproate was amongst the highest ranked drugs (percentile rank 98), but not the highest. The highest ranked prediction was primidone. In the longest retrospective cohort study of JME to date, primidone was most effective, with a 5-year terminal remission rate of 73.3, compared to 50% with valproate.[53]

## For CAE, valproate and ethosuximide are most effective

Valproate and ethosuximide are most effective for CAE; both are similarly effective for CAE.[54] In our predictions for CAE, valproate is ranked highest

of all drugs. Ethosuximide is not ranked highly, but higher than average, amongst our predictions (median percentile rank 58). The *P*-value for the two drugs being ranked so favourably is $5 = 1 \times 10^{-4}$. Ethosuximide is ascribed a particularly low FM score for CAE, which places it in the 20th percentile of predictions for the phenotype. One possible explanation of ethosuximide's low FM score is that its mechanism of action is poorly understood, as it is not an extensively studied compound. Indeed, ethosuximide is one of the least studied of the current ASMs: there are 343 MEDLINE articles with the word ethosuximide in their title, compared to a mean of ~1765 for the other current ASMs that are also found in our datasets (as of 2 September 2021; single-sample one-tailed *t*-test $t = 3.7$ and *P*-value = $6.9 \times 10^{-4}$).

## The drug predictions are not driven by individual highly disease-associated proteins

The relative predicted efficacy of drugs does not change significantly after excluding, one at a time, the top 10 most strongly disease-associated proteins that contribute to the FAM score for that epilepsy. The predicted ranks of drugs for each epilepsy remained significantly stable after excluding, one at a time, the top 10 most strongly disease-associated proteins that contribute to the FAM score for that epilepsy. For each epilepsy, FAM scores were re-calculated after excluding, one at a time, the top 10 most strongly disease-associated proteins (Supplementary Table 3) that contribute to the FAM score for that epilepsy. When drug ranks obtained after excluding a protein were compared with the original drug ranks, Kendall's $\tau$ ranged from 0.80 to 0.93, with all corrected *P*-values $<1 \times 10^{-200}$. In contrast, comparing the predicted drug rankings for two unrelated epilepsies—CAE and HS—yields a Kendall's $\tau$ of 0.04 ($P = 0.10$).

## Top candidate drugs

Ranked lists of the top drugs predicted to be effective for each phenotype, which are most enriched with the drugs that are known to be (more) effective for the phenotype, are available for download (see Data availability). For each phenotype, the top candidate drugs are significantly (Benjamini–Hochberg *P*-value <0.05) enriched with the ASMs that are (more) effective for the

phenotype, except for HS. For HS, there was no significant enrichment of (more) effective ASMs, which may be a reflection of the often drug-resistant nature of HS, or of the lower power of the HS GWAS, or the relatively smaller size of the more effective subset of ASMs for HS, or a combination of these factors.

A manually curated selection of top candidate drugs that could potentially be repurposed for different forms of common epilepsy is shown in the Table 2.



| Epi | Drugs | Evidence of antiseizure efficacy in | Indication | Mode of action |
|---|---|---|---|---|
| CAE | Clomipramine | Animal models[1] and humans[2] | Depression | Serotonin–noradrenaline reuptake inhibitor |
| CAE | Doxepin | Animal models[3,4] | Depression | Tricyclic antidepressant |
| CAE | Pentoxifylline | Animal models[5] | Peripheral vascular disease | Haemorheological agent, increases leukocyte deformability |
| CAE | Phenelzine | Animal models[6] | Depression | Monoamine oxidase inhibitor |
| CAE | Sulindac | Animal models[7] | Pain | Non-steroidal anti-inflammatory |
| CAE | Tolbutamide | Animal models[8] | Diabetes mellitus | Sulphonylurea |
| CAE | Tranylcypromine | Animal models[9] | Depression | Monoamine oxidase inhibitor |
| FE | Chlorzoxazone | Rat hippocampal neurons[10] | Muscle spasms | Calcium and potassium channel inhibitor |
| FE | Hydrochlorothiazide | Animal models[11, 12] and human[12] | Hypertension | ACEII antagonist |
| FE | Thalidomide | Animal models[16–18] | Multiple myeloma | Immunomodulation, unspecified |
| FE | Zaleplon | Animal models[19] | Insomnia | GABA-BZ agonist |
| FE | Zolpidem | Animal models[20–22] | Insomnia | GABA-BZ/GABA-A agonist |
| HS | Amiodarone | Animal models[23] | Arrhythmia | Potassium channel blocker |
| HS | Clonidine | Animal models[24–44] | Hypertension | Alpha-2 adrenoceptor agonist |

| Epi | Drugs | Evidence of antiseizure efficacy in | Indication | Mode of action |
|---|---|---|---|---|
| HS | Methoxamine | Animal models[45] | Hypotension | Alpha-1 adrenergic receptor agonist |
| HS | Pergolide | Animal models[46] | Parkinson's disease | D2 agonist |
| HS | Thioridazine | Animal models[47] | Psychosis | D1/D2 antagonist |
| HS | Tizanidine | Animal models[40] | Muscle spasticity | Alpha-2 adrenergic receptor antagonist |
| JME | Aliskiren | Animal[48, 49] | Hypertension | Renin inhibitor |
| JME | Baclofen | Animal models[39, 50–76] | Muscle spasticity | GABA-B receptor agonist |
| JME | Diazoxide | Animal models[77, 78] | Hypoglycaemia | Potassium channel agonist, inhibits insulin release |
| JME | Icosapent | Animals[79, 80] and humans[81–85] | Hypertriglyceridaemia | 20-carbon omega-3 fatty acid |
| JME | Iloprost | Animal models[86, 87] | Pulmonary arterial hypertension | Synthetic analogue of prostacyclin PGI2 |
| JME | Nicotinamide | Animal models[94–103] | Pellagra | Water-soluble form of Vitamin B3 |
| JME | Pranlukast | Animal models[104] and humans[105] | Asthma | Cysteinyl leukotriene receptor-1 antagonist |
| JME | Riluzole | Animal models[106–109] | Amyotrophic lateral sclerosis | Glutamate antagonist |

**Table 2: Manually curated selection of candidate drugs for the phenotypes shown in the table.** Candidate drugs for GE, which we tested in an animal model, are listed in Table 3. References, for the evidence cited here, can be found in the Supplementary material. CAE, childhood absence epilepsy; Epi, epilepsy type or syndrome; HS, focal epilepsy with hippocampal sclerosis; JME, juvenile myoclonic epilepsy.

## Predicted drugs have a significant dose-dependent effect on seizures in an animal model

After excluding drugs that are toxic or otherwise unsuitable, the top five predicted drugs for GE were tested in a mouse model with a complex genetic seizure disorder that manifests as audiogenic generalized seizures. Each of the drugs had a significant dose-dependent effect on tonic and clonic convulsions (Table 3). Whilst four of the drugs had a significant dose-dependent *anti*-convulsant effect, one of the compounds (betahistine) had a significant dose-dependent *pro*-convulsant effect.

| Drug | Latency (s) to convulsions (mean±s.e.m) | *P* |
|---|---|---|
| Vehicle (i.p.) | 10.9 ± 2.6 | – |
| Orphenadrine (12.5 mg/kg i.p.) | 40.0 ± 5.6 | $6.10 \times 10^{-5}$ |
| Orphenadrine (25 mg/kg i.p.) | 53.4 ± 3.7 | $5.40 \times 10^{-7}$ |
| Orphenadrine (50 mg/kg i.p.) | 60.0 ± 0.0 | $4.14 \times 10^{-7}$ |
| Dyclonine (5 mg/kg i.p.) | 31.5 ± 6.2 | $1.77 \times 10^{-2}$ |
| Dyclonine (10 mg/kg i.p.) | 44.7 ± 5.4 | $2.16 \times 10^{-4}$ |
| Dyclonine (20 mg/kg i.p.) | 57.7 ± 2.4 | $4.14 \times 10^{-7}$ |
| Trimeprazine (2.5 mg/kg i.p.) | 11.0 ± 20.6 | $6.52 \times 10^{-1}$ |
| Trimeprazine (5 mg/kg i.p.) | 18.1 ± 4.1 | $1.77 \times 10^{-2}$ |
| Trimeprazine (10 mg/kg i.p.) | 44.5 ± 5.3 | $4.06 \times 10^{-6}$ |
| Acamprosate (125 mg/kg i.p.) | 8.7 ± 0.4 | $6.40 \times 10^{-1}$ |
| Acamprosate (250 mg/kg i.p.) | 9.2 ± 0.2 | $4.56 \times 10^{-1}$ |
| Acamprosate (500 mg/kg i.p.) | 14.3 ± 2.5 | $1.20 \times 10^{-2}$ |
| Betahistine (75 mg/kg i.p.) | 9.1 ± 0.5 | $4.53 \times 10^{-1}$ |
| Betahistine (150 mg/kg i.p.) | 6.9 ± 0.4 | $2.83 \times 10^{-2}$ |
| Betahistine (300 mg/kg i.p.) | 5.3 ± 0.3 | $4.48 \times 10^{-5}$ |
| Valproate (180 mg/kg i.p.) | 57.7 ± 1.4 | $4.89 \times 10^{-7}$ |

**Table 3: Results from testing compounds in a genetic model of generalised seizures: the DBA/2 mouse model of audiogenic seizures.** After activation of a bell, latency to the occurrence of tonic convulsions and clonic convulsions was measured. P, Benjamini–Hochberg–corrected P-value from two-sided Mann–Whitney U test; s.e.m, standard error of the mean.



# Discussion

We present the relative predicted efficacy of drugs against each of the main types and syndromes of common epilepsy. This dataset is a novel and valuable resource for selecting the best candidate drug(s) to repurpose for any of the main types and syndromes of common epilepsy. Of course, our predicted candidate drugs require further animal model and/or human clinical trial evidence before being considered for deployment in clinical practice.

To generate our predictions, we created a novel method. Our method possesses several strengths that are lacking in previously published

approaches. Common epilepsies, like other complex diseases, develop when many different proteins display abnormal activity due to pathological changes in their abundance or function.[55] Our method prioritizes drugs according to their relative ability to modulate changes in both the abundance and the function of disease-proteins. Furthermore, drugs are prioritized on the basis of their ability to correct disease-protein abnormalities that are found in people with the disease, rather than in animal models, and that are not consequential to or compensatory for the disease, as they are driven by germline variations. We use genetic variation data specific to each form of common epilepsy, to make drug predictions specific to that form of common epilepsy. The ASMs that are more clinically-effective for a syndrome and the ASMs that are less clinically-effective for a syndrome are predicted more effective and less effective, respectively, for that syndrome only, but not for any other epilepsy type or syndrome—this suggests that our predictions are not systemically biased in favour of a particular set or type of drugs. The methodology is based upon a polygenic model of disease and a multi-targeted approach to treatment, which are desirable for complex diseases. We utilize conventional canonical low-throughput single-target functional drug activity data, and high-throughput genome-wide transcriptomic drug activity data, so that prioritization of drugs is informed by their on-target and off-target effects, and by their affinities for individual proteins and effects upon genome-wide gene expression. The directionality of drugs' effects on protein activity also helps inform drug prioritization. Rather than dichotomous categorization of compounds into drugs that are predicted to be effective or ineffective, our method ranks drugs individually according to relative predicted efficacy, which aids candidate selection for *in vivo* validation and for development.

Our method produces accurate drug predictions for epilepsy syndromes even if their GWAS results include few genome-wide significant loci. Even excluding the most strongly disease-associated proteins does not significantly change the relative efficacy of drugs predicted by our method (as we show in the Results, under subheading 'The drug predictions are not driven by individual highly disease-associated proteins'). This is because our method is not reliant on individual highly disease-associated proteins.

Instead, our method leverages the gene-set analysis approach, where each gene-set is the set of genes affected by each drug. The disease association of all the genes in a gene-set, even those below the genome-wide significant threshold, is combined; the gene-sets that are more disease-associated overall are more biologically relevant. The gene-set approach is a long-established and widely-used method in all areas of genomic analysis,[56] including post-GWAS analysis generally[57] and GWAS-based drug repurposing analysis specifically.[27,28] Utilizing the full distribution of all genetic associations for gene-set analysis is a validated, established and accepted approach, which has been implemented in numerous widely-used post-GWAS analysis tools, for example FUMA,[58] MAGMA,[58] MAGENTA,[59] INRICH[60] and DEPICT,[61] each of which has been employed in a multitude of published GWAS-based studies.



Alongside these strengths, our method has some limitations, discussed below.

Our drug prediction method, like all previously published genetics- or genomics-based drug prediction methods, predicts the efficacy of drugs for a disease. However, the most efficacious drug for a disease in not always the most appropriate drug for every individual with the disease. Important factors to consider when choosing a drug for an individual include the potential of undesirable interactions with other medications and the possible side-effects. Our method, like all previously published genetics- or genomics-based drug prediction methods, does not predict drugs' interactions with other medications and side-effects. Indeed, the success of an ASM is determined as much by its tolerability as by its efficacy.[59–62] As the drugs we have predicted for epilepsies are already being used for other diseases, their side-effect profiles are known. This allows researchers to select for further development those candidate compounds whose side-effects are less deleterious or even desirable.

Our method predicts drugs effective for a disease from the proteins underlying the disease, after identifying the proteins underlying the disease from the common genetic variations associated with the disease. However, some proteins become dysfunctional or dysregulated not because of

common genetic variations, but because of rare genetic variations, or copy number variations, or abnormalities of epigenetic, post-transcriptional or post-translational mechanisms, or because of environmental insults. Such protein changes do not inform our predictions, which could affect their accuracy, commensurate with the contribution of those proteins to the causal mechanism underlying an epilepsy and/or to the mechanism of action of a drug. We are not aware of any existing drug prediction methods which take into account the multiple potential pathogenic factors that influence proteins; the development of such methods might lead to improved accuracy of drug predictions.

Our analysis uses data from a GWAS that employed imputation to improve genomic coverage. The GWAS gene-level data used in this analysis offers coverage of genes across the genome, and it is corrected for the lengths and single nucleotide polymorphism-densities of genes. However, if a gene is not (adequately) covered by the genotyping array and the imputation, but the gene is of importance in epilepsy and affected by drug(s), the accuracy of our drug predictions could be adversely affected. Hence, improved coverage of future epilepsy GWAS analyses could help to improve the accuracy of drug predictions.

Our drug predictions are based upon two scores: the FM and AC scores. The FM score relies upon knowledge of the proteins changed in function by drugs. At present, knowledge of the proteins that are changed in function by each drug is incomplete, and it is more incomplete for some drugs than for others. The more incomplete the knowledge of the proteins changed in function by a drug, the more likely it is that the drug's FM score will be underestimated. By extension, the FM score is more likely to be underestimated for drugs that are less studied, as their modes of action are less analysed and, hence, knowledge of the proteins changed in function by them is less complete. This may explain the relatively low FM sore and, hence, FAM score and ranking for ethosuximide. The AC score is free of this limitation, as the AC score is based upon profiles of drug-induced transcriptomic changes assayed by using the same standardized pipeline for each drug. With over 44 000 compounds already analysed on this platform (http://lincsportal.ccs.miami.edu/SmallMolecules/catalog; accessed on 1

February 2021), transcriptomic profiles are available for the vast majority of drugs of interest. However, a small number of interesting drugs (for example, brivaracetam and cenobamate) have not been assayed, which means that an AC score and, hence, a FAM score cannot be calculated for them. The platform and pipeline used for generating drugs' transcriptomic profiles are in the public domain, and have been used by researchers to generate profiles for any compounds of interest not already found in the database, albeit for industrial-scale projects.[62] In addition, there is active ongoing development of computational methods for using knowledge of drugs' structures to predict the proteins that they change in function and/or abundance,[63,64] which is another potential future strategy for predicting the relative efficacy of compounds whose molecular effects are still unknown.



It is noted that the FM score does not predict the 'directionality' of drugs' effects (that is, beneficial or harmful) on disease-protein function. Therefore, drugs predicted by the FM score to affect a phenotype may be alleviating or aggravating for the phenotype. This is a recognized limitation of methods that use data for the ability of drugs to alter the function of genetically-associated disease-proteins in order to predict drugs that can affect the disease,[16,17,65] as the direction of change in protein activity occurring in the disease is unknown. On the other hand, the AC score does predict the 'directionality' of drugs' effects (that is, beneficial or harmful) on individual disease-proteins and, thereby, the overall 'directionality' of drugs' effects (that is, beneficial or harmful) on the disease. The AC score takes into account the magnitude and direction of change in proteins' abundance underlying disease, and the magnitude and direction of change in proteins' abundance caused by drugs. Thereby, the AC score proposes to predict the drugs with a beneficial effect on disease-protein abundance and clinical phenotype. Hence, inclusion of the AC score, with the FM score, in our final FAM score, is expected to help mitigate the risk of deleterious compounds with high FM scores being included in our candidate drugs. Still, it is possible that some aggravating drugs are included in our candidate compounds. Hence, experimental validation of candidate drugs is essential before clinical use, as with all *in silico* drug prediction methods. We tested five of our candidate compounds in a rodent model: all five compounds

had a significant dose-dependent effect on seizures. Interestingly, one of the candidate compounds (betahistine) had a significant dose-dependent pro-convulsant effect in the animal model. This finding could be explained by the possibility that some of our predicted compounds are aggravating, as discussed. However, it is also possible that the pro-convulsant effect of betahistine in our study is a reflection of species- or model-specific behaviour. Indeed, a recent study (published after our animal experiments had ended) showed that betahistine has a significant antiepileptogenic and anticonvulsant effect on pentylenetetrazole-induced generalized seizures in a different mouse strain.[66]

Whilst acknowledging these limitations and some aberrant predictions, we note that our method outperforms alternative methods for predicting drugs that have efficacy against common epilepsies in clinical studies and experience. Our method also predicts which ASMs are amongst the more efficacious in clinical practice, and which ASMs are amongst the less efficacious in clinical practice, for each of the main syndromes of common epilepsy, and it predicts the distinct order of efficacy of individual ASMs in clinical trials of different common epilepsies. This aspect is key to the clinical translation of drug predictions for common epilepsies, but is missing from previously published studies that have predicted drugs for epilepsy.[13–17]

In this study, we have used the tissue-wide association study method to identify the protein abundance changes underlying disease. A closely-related alternative method is to use Mendelian randomization. In future studies, both methods could be compared and/or combined in order to determine if this improves the drug predictions. Mendelian randomization is discussed at greater length in the Supplementary material.

As our method uses GWAS data, it cannot be applied to monogenic diseases. It is conceivable that this method could be adapted to make it applicable to monogenic diseases, and we plan to explore this possibility in a future study dedicated to this objective.

We have used results from the latest epilepsy GWAS mega-analysis, which includes previously published and unpublished epilepsy GWAS analyses,

making it the largest epilepsy GWAS to date.[11] However, compared to other common neurological diseases, even the largest epilepsy GWAS had a modest sample size, with 15 212 cases and 29 677 controls, and produced a modest number of discoveries, with 16 loci identified. The latest schizophrenia GWAS, for example, included 36 989 cases and 113 075 controls, resulting in the identification of 108 risk loci.[67] It is hoped that expanded cohort sizes of future epilepsy GWAS analyses will increase power and improve drug predictions. In this analysis, we predicted drugs for the main epilepsy syndromes that had risk loci identified in the latest epilepsy GWAS. It is hoped that future epilepsy GWAS will be large enough to report results for additional epilepsy syndromes, and drugs can be predicted for them using the method presented here. Finally, it is likely that our method can be applied to the GWAS results of other common complex phenotypes.



## Supplementary material

Supplementary material is available at: https://tinyurl.com/4j3srm5j.

## Acknowledgements

Some of the data reported in this paper were collected as part of a project undertaken by the International League against Epilepsy (ILAE) and some of the authors are experts selected by the ILAE. Opinions expressed by the authors, however, do not necessarily represent the policy or position of the ILAE.

## Funding

This work was supported by a Medical Research Council Confidence in Concept Award to N.M. B.P.C.K. and R.S are supported by funds from the Friends of Wilhemina KinderZiekenhuis MING funds.

## Competing interests

The authors report no competing interests.

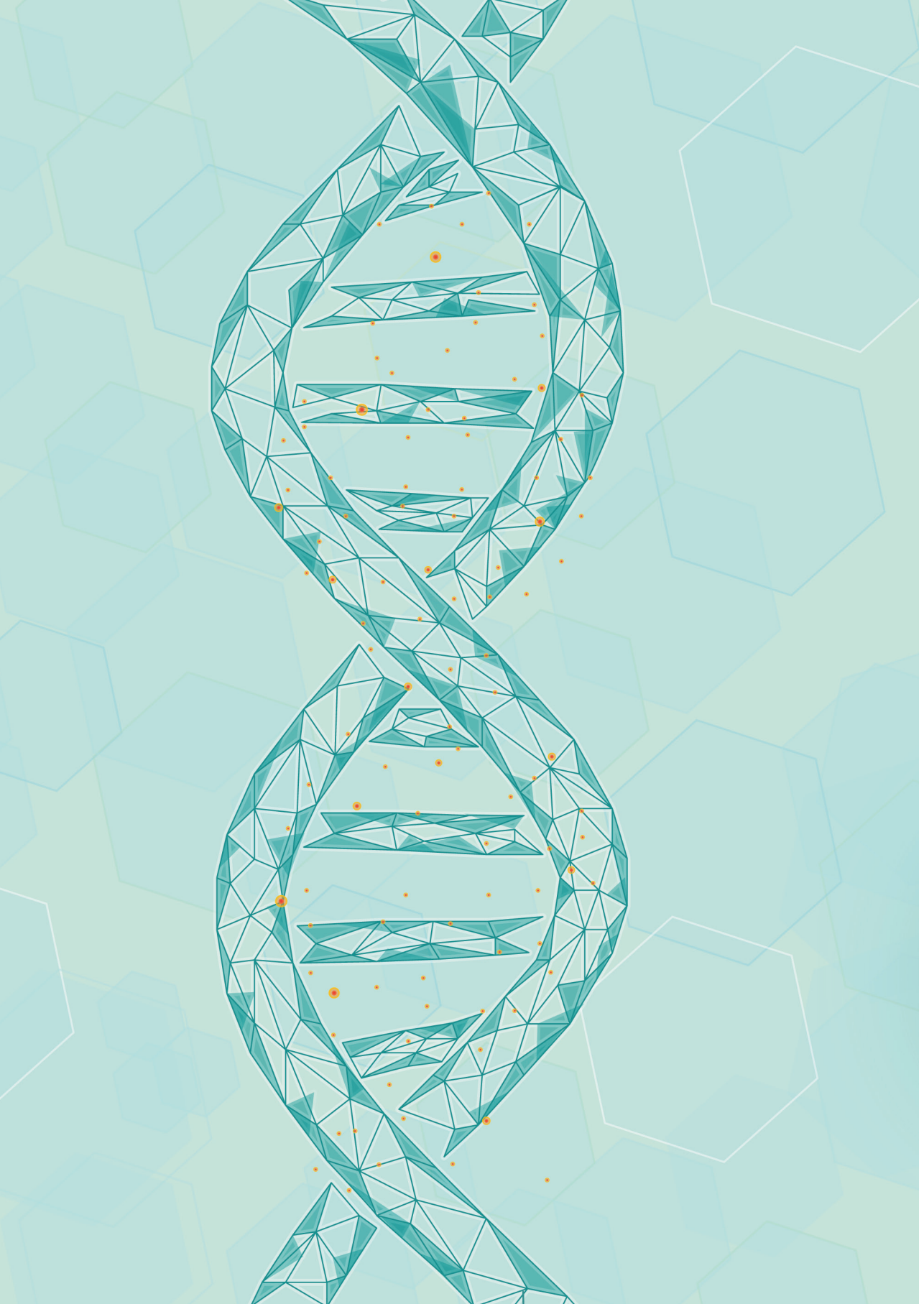
# References


1. GBD 2016 Epilepsy Collaborators. Global, regional, and national burden of epilepsy, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2019;18(4):357–375.
2. Shorvon SD. The epidemiology and treatment of chronic and refractory epilepsy. Epilepsia. 1996;37 Suppl 2:S1–S3.
3. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: A 30-year longitudinal cohort study. JAMA Neurol. 2018;75(3):279–286.
4. Baker GA, Jacoby A, Buck D, Stalgis C, Monnet D. Quality of life of people with epilepsy: A European study. Epilepsia. 1997;38(3):353–362.
5. Perucca P, Carter J, Vahle V, Gilliam FG. Adverse antiepileptic drug effects: Toward a clinically and neurobiologically relevant taxonomy. Neurology. 2009;72(14):1223–1229.
6. Alsfouk BAA, Brodie MJ, Walters M, Kwan P, Chen Z. Tolerability of antiseizure medications in individuals with newly diagnosed epilepsy. JAMA Neurol. 2020;77(5):574–581.
7. Sivapalarajah S, Krishnakumar M, Bickerstaffe H, et al. The prescribable drugs with efficacy in experimental epilepsies (PDE3) database for drug repurposing research in epilepsy. Epilepsia. 2018;59(2):492–501.
8. Kamb A, Harper S, Stefansson K. Human genetics as a foundation for innovative drug development. Nat Biotechnol. 2013;31(11):975–978.
9. Koeleman BPC. What do genetic studies tell us about the heritable basis of common epilepsy? Polygenic or complex epilepsy? Neurosci Lett. 2018;667:10–16.
10. Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017;58(4):512–521.
11. 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(1):5269.
12. Bourgeois BF. Chronic management of seizures in the syndromes of idiopathic generalized epilepsy. Epilepsia. 2003;44 Suppl 2:27–32.
13. Mirza N, Sills GJ, Pirmohamed M, Marson AG. Identifying new antiepileptic drugs through genomics-based drug repurposing. Hum Mol Genet. 2017;26(3):ddw410.
14. Brueggeman L, Sturgeon ML, Martin RM, et al. Drug repositioning in epilepsy reveals novel antiseizure candidates. Ann Clin Transl Neurol. 2019;6(2):295–309.
15. Delahaye-Duriez A, Srivastava P, Shkura K, et al. Rare and common epilepsies converge on a shared gene regulatory network providing opportunities for novel antiepileptic drug discovery. Genome Biol. 2016;17(1):245.
16. Himmelstein DS, Lizee A, Hessler C, et al. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. Elife. 2017;6:e26726.
17. Guney E, Menche J, Vidal M, Barabasi AL. Network-based in silico drug efficacy screening. Nat Commun. 2016;7:10331.

18. Steffens M, Leu C, Ruppert A-K, et al. ; EMINet Consortium. Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. Hum Mol Genet. 2012;21(24):5359–5372.
19. International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address: epilepsy-austin@unimelb.edu.au. Genetic determinants of common epilepsies: A meta-analysis of genome-wide association studies. Lancet Neurol. 2014;13(9):893–903.
20. Kasperaviciute D, Catarino CB, Matarin M, et al. ; UK Brain Expression Consortium. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. Brain. 2013;136(Pt 10):3140–3150.
21. Sanseau P, Agarwal P, Barnes MR, et al. Use of genome-wide association studies for drug repositioning. Nat Biotechnol. 2012;30(4):317–320.
22. Young MP, Zimmer S, Whitmore AV, Morphy J, Harris C. Drug molecules and biology: Network and systems aspects. Cambridge, UK: Designing multi-target drugs Royal Society of Chemistry; 2012:32–49.
23. Iskar M, Campillos M, Kuhn M, Jensen LJ, van Noort V, Bork P. Drug-induced regulation of target expression. PLoS Comput Biol. 2010;6(9):e1000925.
24. Bantscheff M, Scholten A, Heck AJ. Revealing promiscuous drug-target interactions by chemical proteomics. Drug Discov Today. 2009;14(21–22):1021–1029.
25. So HC, Chau CK, Lau A, Wong SY, Zhao K. Translating GWAS findings into therapies for depression and anxiety disorders: Gene-set analyses reveal enrichment of psychiatric drug classes and implications for drug repositioning. Psychol Med. 2019;49(16):2692–2708.
26. So HC, Chau CK, Chiu WT, et al. Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. Nat Neurosci. 2017;20(10):1342–1349.
27. Lau A, So HC. Turning genome-wide association study findings into opportunities for drug repositioning. Comput Struct Biotechnol J. 2020;18:1639–1650.
28. Reay WR, Cairns MJ. Advancing the use of genome-wide association studies for drug repurposing. Nat Rev Genet. 2021;22(10):658–671.
29. Gallagher MD, Chen-Plotkin AS. The post-GWAS era: From association to function. Am J Hum Genet. 2018;102(5):717–730.
30. Sirota M, Dudley JT, Kim J, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med. 2011;3(96):96ra77.
31. Cheng J, Xie Q, Kumar V, et al. Evaluation of analytical methods for connectivity map data. Pac Symp Biocomput. 2013;5–16.
32. Okada Y, Wu D, Trynka G, et al. ; GARNET Consortium. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014;506(7488):376–381.
33. Wang YF, Zhang Y, Zhu Z, et al. Identification of ST3AGL4, MFHAS1, CSNK2A2 and CD226 as loci associated with systemic lupus erythematosus (SLE) and evaluation of SLE genetics in drug repositioning. Ann Rheum Dis. 2018;77(7):1078–1084.
34. Wain LV, Shrine N, Artigas MS, et al. ; Geisinger-Regeneron DiscovEHR Collaboration. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. Nat Genet. 2017;49(3):416–425.

35. Aguirre-Plans J, Pinero J, Sanz F, et al. GUILDify v2.0: A tool to identify molecular networks underlying human diseases, their comorbidities and their druggable targets. J Mol Biol. 2019;431(13):2477–2484.
36. So HC, Chau CK, Lau A, Wong SY, Zhao K. Translating GWAS findings into therapies for depression and anxiety disorders: Gene-set analyses reveal enrichment of psychiatric drug classes and implications for drug repositioning. Psychol Med. 2019;49:2692–2708.
37. Grau J, Grosse I, Keilwagen J. PRROC: Computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics. 2015;31(15):2595–2597.
38. Abdi H. The Kendall rank correlation coefficient. Thousand Oaks, CA: Encyclopedia of Measurement and Statistics Sage; 2007:508–510.
39. Misawa H, Sherr EH, Lee DJ, et al. Identification of a monogenic locus (jams1) causing juvenile audiogenic seizures in mice. J Neurosci. 2002;22(23):10088–10093.
40. Italiano D, Striano P, Russo E, et al. Genetics of reflex seizures and epilepsies in humans and animals. Epilepsy Res. 2016;121:47–54.
41. Frankel WN. Genetics of complex neurological disease: Challenges and opportunities for modeling epilepsy in mice and rats. Trends Genet. 2009;25(8):361–367.
42. Neumann PE, Seyfried TN. Mapping of two genes that influence susceptibility to audiogenic seizures in crosses of C57BL/6J and DBA/2J mice. Behav Genet. 1990;20(2):307–323.
43. Dürmüller N, Smith SE, Meldrum BS. Proconvulsant and anticonvulsant effects of Evans blue dye in rodents. Neuroreport. 1993;4(6):683–686.
44. Marson AG, Sills GJ. Valproate. In: Shorvon, S., Perucca, E. and Engel Jr, J. eds. The treatment of epilepsy. John Wiley & Sons; 2015:652–666.
45. Tomson T, Johannessen SI. Carbamazepine. In: Shorvon, S., Perucca, E. and Engel Jr, J. eds. The treatment of epilepsy. John Wiley & Sons; 2015:431–446.
46. Marson AG, Al-Kharusi AM, Alwaidh M, et al.; SANAD Study Group. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: An unblinded randomised controlled trial. Lancet. 2007;369(9566):1000–1015.
47. Marson A, Burnside G, Appleton R, et al.; SANAD II Orators. The SANAD II study of the effectiveness and cost-effectiveness of levetiracetam, zonisamide, or lamotrigine for newly diagnosed focal epilepsy: An open-label, non-inferiority, multicentre, phase 4, randomised controlled trial. Lancet. 2021;397(10282):1363–1374.
48. Marson AG, Al-Kharusi AM, Alwaidh M, et al.; SANAD Study Group. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: An unblinded randomised controlled trial. Lancet. 2007;369(9566):1016–1026.
49. Marson A, Burnside G, Appleton R, et al. ; SANAD II Orators. The SANAD II study of the effectiveness and cost-effectiveness of valproate versus levetiracetam for newly diagnosed generalised and unclassifiable epilepsy: An open-label, non-inferiority, multicentre, phase 4, randomised controlled trial. Lancet. 2021;397(10282):1375–1386.

50. Brodie MJ. Modern management of juvenile myoclonic epilepsy. Expert Rev Neurother. 2016;16(6):681–688.
51. Crespel A, Gelisse P, Reed RC, et al. Management of juvenile myoclonic epilepsy. Epilepsy Behav. 2013;28 Suppl 1:S81–S86.
52. Nicolson A, Marson AG. When the first antiepileptic drug fails in a patient with juvenile myoclonic epilepsy. Pract Neurol. 2010;10(4):208–218.
53. Senf P, Schmitz B, Holtkamp M, Janz D. Prognosis of juvenile myoclonic epilepsy 45 years after onset: Seizure outcome and predictors. Neurology. 2013;81(24):2128–2133.
54. Glauser TA, Cnaan A, Shinnar S, et al. Ethosuximide, valproic acid, and lamotrigine in childhood absence epilepsy. N Engl J Med. 2010;362(9):790–799.
55. Speed D, O'Brien TJ, Palotie A, et al. Describing the genetic architecture of epilepsy through heritability analysis. Brain. 2014;137(Pt 10):2680–2689.
56. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–15550.
57. Luo L, Peng G, Zhu Y, Dong H, Amos CI, Xiong M. Genome-wide gene and pathway analysis. Eur J Hum Genet. 2010;18(9):1045–1053.
58. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
59. Segre AV; DIAGRAM Consortium, MAGIC Investigators, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 2010;6(8):e1001058.
60. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: Interval-based enrichment analysis for genome-wide association studies. Bioinformatics. 2012;28(13):1797–1799.
61. Pers TH, Karjalainen JM, Chan Y, et al. ; Genetic Investigation of ANthropometric Traits (GIANT) Consortium. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6:5890.
62. De Wolf H, Cougnaud L, Van Hoorde K, et al. High-throughput gene expression profiles to define drug similarity and predict compound activity. Assay Drug Dev Technol. 2018;16(3):162–176.
63. Hu S, Zhang C, Chen P, Gu P, Zhang J, Wang B. Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks. BMC Bioinformatics. 2019;20(Suppl 25):689.
64. Hodos R, Zhang P, Lee HC, et al. Cell-specific prediction and application of drug-induced gene expression profiles. Pac Symp Biocomput. 2018;23:32–43.
65. Cheng F, Desai RJ, Handy DE, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. Nat Commun. 2018;9(1):2691.
66. Yazdi A, Doostmohammadi M, Pourhossein Majarshin F, Beheshti S. Betahistine, prevents kindling, ameliorates the behavioral comorbidities and neurodegeneration induced by pentylenetetrazole. Epilepsy Behav. 2020;105:106956.
67. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421–427.



## CHAPTER 6

# GENOME-WIDE META-ANALYSIS OF OVER OF 29,000 PEOPLE WITH EPILEPSY REVEALS 26 LOCI AND SUBTYPE-SPECIFIC GENETIC ARCHITECTURE

International League Against Epilepsy Consortium on Complex Epilepsies
Contribution: lead author and analist

*Manuscript under review*

## Abstract

Epilepsy is a highly heritable disorder affecting around 70 million people worldwide. Approximately one third of people with epilepsy experience seizures refractory to current treatments. Much of the heritability is attributable to common variants, suggesting that genome-wide association studies (GWAS) can yield biological insights and aid pathophysiology-informed drug discovery. Here, we report a trans-ethnic GWAS including 29,944 specifically phenotyped cases and 52,538 controls, stratified into three broad- and seven subtypes of epilepsy. We further increased the sample size using biobank-derived individuals to 51,678 cases and 1,076,527 controls. We identify 26 genome-wide significant loci. Nineteen of these signals are specific to genetic generalised epilepsy (GGE) and three to the GGE subtype juvenile myoclonic epilepsy, whilst three are pleiotropic and one is female-specific. We implicate 29 likely causal genes underlying these 26 loci, based on a combination of ten prioritization methods. SNP-based heritability analyses show that common variants largely close the gap in missing heritability for GGE. Subtype analysis revealed markedly different genetic architectures between focal and generalised forms of epilepsy. Enrichment analyses implicate synaptic processes in excitatory as well as inhibitory neurons in the brain. Genes identified in our study overlap with monogenic epilepsy genes and targets of current anti-epileptic drugs. Finally, we leverage our GWAS to highlight a list of drugs with predicted efficacy when repurposed for epilepsy treatment.

# Introduction

The epilepsies are a group of heterogeneous neurological disorders, characterized by an enduring predisposition to generate unprovoked seizures.[1] It is estimated that over 50 million people worldwide have active epilepsy, with an annual cumulative incidence of 67.7 per 100,000 persons.[2]

Similar to other common neurodevelopmental disorders, the epilepsies have substantial genetic risk contributions from both common and rare genetic variation. Analysis of the epilepsies benefits from well characterized phenotyping which allows clinical sub groups to be distinguished, in contrast to other neurodevelopmental disorders where phenotypic subgroups are more difficult to define. Differences in the genetic architecture of these clinical subgroups of epilepsies are also emerging to complement the clinical partitioning.[3–5] The rare but severe epileptic encephalopathies are usually non-familial and are largely caused by a heterogeneous collection of *de novo* dominant variants, often involving ion channels or synaptic machinery.[6] Common, as well as rare variation, has been shown to contribute to the milder and more common focal and generalized epilepsies, both of which have high heritability. This is particular generalized epilepsy, which is primarily constituted by genetic generalised epilepsy (GGE).[3,4,7,8] Nevertheless, previous genetic studies of common epilepsies have explained only a few percent of this common genetic, or SNP-based, heritability.[3–5,8]

Epilepsy is typically treated using antiepileptic drugs (AEDs). However, despite the availability of over 25 licensed AEDs worldwide, a third of people with epilepsy experience ongoing seizures.[9] Diet, surgery and neuromodulation represent additional treatment options that can be effective in small subgroups of patients.[10] Accurate classification of patients is an important guiding factor in epilepsy treatment.

Here, we report the third epilepsy GWAS meta-analysis, comprising a total of 29,944 specifically phenotyped cases recruited from tertiary referral centres, and 52,538 controls, approximately doubling the previous sample size.[3] Results suggest markedly different genetic architectures between focal and generalised forms of epilepsy. Combining these results with results from less stringently phenotyped biobank-derived epilepsy did not substantially

increase signal, despite almost doubling the sample size to 51,678 cases and 1,076,527 controls. Our findings shed light on the enigmatic biology of generalized epilepsy, the importance of accurate syndromic phenotyping and may facilitate drug repurposing for novel therapeutic approaches.

# Results

## Study overview

We performed a genome-wide meta-analysis by combining the previously published effort from our consortium[3] with unpublished data from the Epi25 collaborative[8] and four additional cohorts (**Supplementary table 1**). Our primary mixed model meta-analysis constitutes 4.9 million SNPs tested in 29,944 people with epilepsy and 52,538 controls, of which 16,447 people had focal epilepsy and 7,407 people had GGE. The epilepsy cases were primarily of European descent (92%), with a smaller proportion of African (3%) and Asian (5%) ancestry. Cases were matched with controls of the same ancestry and GWAS were performed separately per ancestry, before performing trans-ethnic meta-analyses. We performed meta-analyses for the broad epilepsy phenotypes 'focal epilepsy' (n=16,447 cases) and 'GGE' (n=7,407 cases). We further conducted meta-analysis of the well-defined GGE subtypes of: a) juvenile myoclonic epilepsy (JME), b) childhood absence epilepsy (CAE), c) juvenile absence epilepsy (JAE) and d) generalized tonic-clonic seizures only (GTCSA), as well as the focal epilepsy subtypes of: a) focal epilepsy with hippocampal sclerosis,b) focal epilepsy due to other lesions, and c) lesion-negative focal epilepsy. We ran a variety of downstream analyses to identify potential sex-specific signals and obtain biological insights and leads for drug-repurposing.

## GWAS for the epilepsies

Our 'all epilepsy' meta-analysis revealed four genome-wide significant loci, of which two were novel (**Figure 1**). Similar to our previous GWAS, the 2q24.3 locus was composed of two independently significant signals.[3] Furthermore, a novel suggestive signal (rs4932477, $p=5.04*10^{-8}$) was found at chromosome 15, containing *POLG*, which is associated with one of the most severe kinds of intractable monogenic epilepsy.[11] Using ASSET to determine the extent of FE and GGE-related pleiotropy, the 2q24 and 9q21 signals showed pleiotropic

effects at a genome-wide significance level, with concordant SNP effect directions for both forms of epilepsy (**Supplementary table 2**). The 2p16.1 and 10q24.32 loci were primarily driven by GGE. The 'focal epilepsy' (FE) analysis did not reveal any genome-wide significant signals.

Analysis of 'generalized epilepsy' (GGE) uncovered a total of 25 independent genome-wide significant signals across 22 loci, of which 14 loci are novel. The strongest signal of association ($p=6.58E^{-21}$), located at 2p16.1, constitutes three independently significant signals. Similarly, the novel locus 12q13.13 was composed of two independently significant signals.

Functional annotation of the 2,355 genome-wide significant SNPs across the 22 loci revealed that most variants were intergenic or intronic (**Supplementary data 1**). 26/2355 (1.1%) SNPs were exonic, of which 12 were located in protein-coding genes and nine were missense variants. Sixty one percent of SNPs were located in open chromatin regions, as indicated by a minimum chromatin state of 1-7.[12] Further annotation by Combined Annotation-Dependent Depletion (CADD) scores predicted 110 associating SNPs to be deleterious (CADD score >12.37).[13] LDAK heritability analyses showed significant enrichment of signal in super-enhancers (**Supplementary table 3**), suggesting that GGE variants regulate expression of genes that define cell identity.[14]

To assess potential syndrome-specific loci, we performed GWAS on seven well-defined focal and GGE subtypes **(Supplementary figure 1A-G)**. We found three genome-wide significant loci associated with JME, of which one was novel (8q23.1), and the other two (4p12 and 16p11.2) reported in our previous GWAS.[3] All three signals appear specific to JME; they were not even nominally associated with any other GGE subtype and did not reach genome-wide significance in the combined GGE analysis. Our analysis of CAE consolidated an established genome-wide significant signal at 2p16.1, which was also observed in the GGE and all epilepsy GWAS. We did not find any genome-wide significant loci for JAE, GTCS, 'non-lesional focal epilepsy', 'focal epilepsy with hippocampal sclerosis', or' focal epilepsy due to other lesions'.

Genomic inflation factors were comparable to our previous GWAS and all linkage-disequilibrium score regression (LDSR) intercepts were lower than our previous GWAS (**Supplementary table 4**),[3] suggesting that the

signals are primarily driven by polygenicity, rather than by confounding or population stratification.[15]

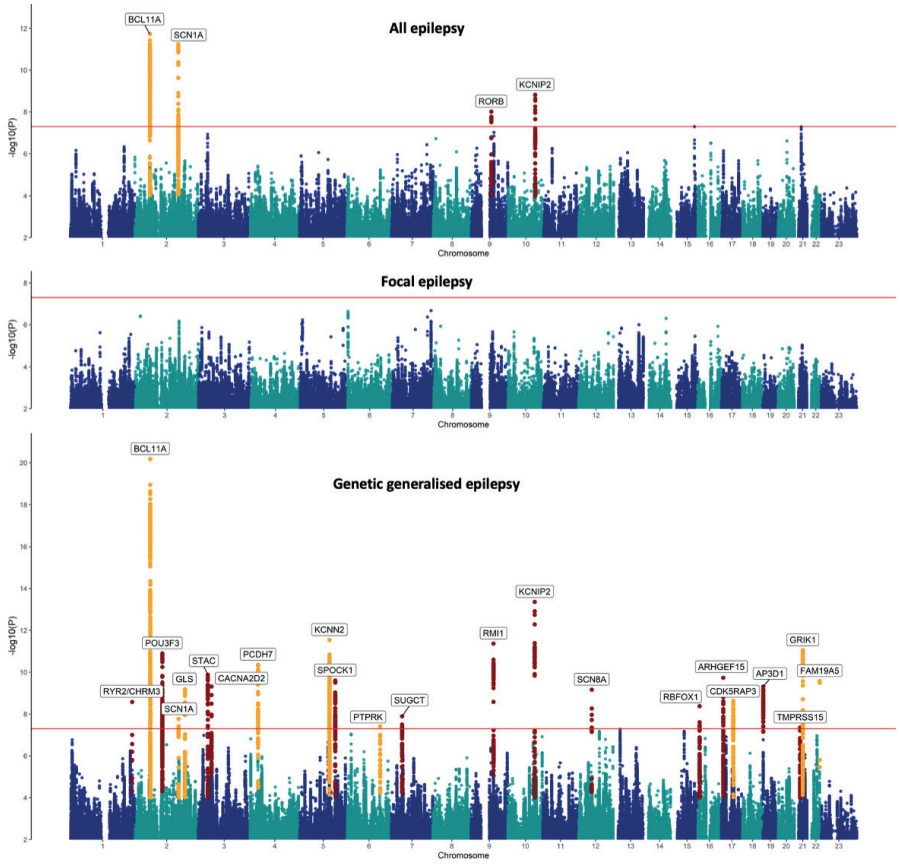



**Figure 1. Manhattan plot of trans-ethnic all, focal epilepsy and genetic generalised epilepsy (GGE) genome-wide meta-analyses.** The red line shows the genome-wide significance threshold ($5x10^{-8}$). Chromosome and position is displayed on the x axis and -log10 P-value on the y axis. Novel genome-wide significant loci are highlighted in red and replicated loci are labeled in orange. Annotated genes are those implicated by our gene prioritization analyses.

## Locus annotation, TWAS and gene prioritisation

Using FUMA[16] (see Methods), the 'all epilepsy' meta-analysis was mapped to 43 genes and the GGE analysis to 278 genes (**Supplementary data 2**). Thirty nine of the 43 'all epilepsy' genes overlapped with GGE, resulting in a total of 282 uniquely mapped genes. These 282 genes were enriched for monogenic

epilepsy genes (hypergeometric test, 18/837 genes overlapped; odds ratio [OR]=1.51, P=0.041), and targets of anti-epileptic drugs (hypergeometric test, 9/191 genes overlap; OR=3.39, P=5.4e-4).

We calculated a gene-based association score based on the aggregate of all SNPs inside each gene using MAGMA (see Methods).[17] This analysis yielded 39 significant genic associations, six with 'all epilepsy', and 37 with GGE (four overlapped with the 'all epilepsy' analysis), after correction for 16,371 tested genes ($p<0.05/16,371$ genes; **Supplementary data 3**). Thirteen of these 39 genes mapped to regions outside of the genome-wide significant loci from the single SNP analyses.

Next, we performed a transcriptome-wide association study (TWAS) to assess whether epilepsy was associated with differential gene expression in the brain (see Methods).[18,19] These analyses revealed that expression of 13 and 16 genes, significantly implicated with 'all epilepsy' and GGE respectively (**Supplementary data 4**). 19 of these genes mapped outside of the 26 loci identified through the GWAS. Using Summary-data-based Mendelian Randomization (SMR)[20], we determined a potentially causal relationship between brain expression of *RMI1* and 'all epilepsy', and between *RMI1*, *CDK5RAP3*, *TVP23B* and GGE (**Supplementary data 5**).

Of note, expression of *RMI1* was associated with GGE in both TWAS (p=4.01E-10) and SMR (p=5.21E-08), as well as with 'all epilepsy' (TWAS p=1.32E-06; SMR p=2.29E-06). *RMI1* has a crucial role in genomic stability[21] and has not been previously associated with epilepsy nor any other Mendelian trait (OMIM #610404).

We used a combination of ten different criteria to identify the most likely implicated gene within each of the 26 associated loci from the meta-analysis (see Methods). This resulted in a shortlist of 29 genes (**Figure 2**), of which ten are monogenic epilepsy genes, seven are known targets of currently licensed anti-epileptic drugs and 17 are associated with epilepsy for the first time. Interrogation of the Drug Gene interaction database (DGIdb) showed that 13 of the 29 genes are targeted by a total of 214 currently licensed drugs (**Supplementary data 6**).

The strongest association signal for GGE was found at 2p16.1, consistent with our previous results where we implicated the gene *VRK2* or *FANCL*.[22] Our gene

prioritization analysis now points to the transcription factor *BCL11A* as the culprit gene, located 2.5MB upstream of the lead SNPs at this locus. Two of three lead SNPs are located in enhancer regions (as assessed by chromatin states in brain tissue) which are linked to the *BCL11A* promoter via 3D chromatin interactions (**Supplementary figure 2**). Rare variants in *BCL11A* were recently associated with intellectual disability and epileptic encephalopathy[23]. However, interrogation of the MetaBrain eQTL database did not reveal a significant association between our lead SNPs with *BCL11A* expression.

| Phenotype | Locus | Novel/replication | Lead SNP (A1:A2) | Freq1 | Z-score | P-value | Gene | Total | Missense | TWAS | SMR | MAGMA | PoPS | Brain exp | brain-coX | KO mouse | AED target | Monogenic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| All epilepsy | 2p16.1 | Replication | rs13032423 (A:G) | 0.53 | -7.04 | 1.85E-12 | *BCL11A* | 5 | | | | | ■ | ■ | ■ | ■ | | ■ |
| | 2q24.3 | Replication | rs59237858 (T:C) | 0.23 | -6.89 | 5.75E-12 | *SCN1A* | 8 | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| | 9q21.13 | Novel | rs4744696 (A:G) | 0.82 | -5.74 | 9.69E-09 | *RORB* | 4 | | | | | ■ | ■ | ■ | ■ | | |
| | 10q24.32 | Novel | rs3740422 (C:G) | 0.33 | 6.04 | 1.52E-09 | *KCNIP2* | 3 | | | | ■ | | ■ | ■ | | | |
| GGE | 1q43 | Novel | rs876793 (T:C) | 0.67 | -5.95 | 2.64E-09 | *RYR2* | 4 | | | | | ■ | ■ | ■ | ■ | | |
| | | | | | | | *CHRM3* | 4 | | | | | | ■ | ■ | ■ | ■ | ■ |
| | 2p16.1 | Replication | rs11688767 (A:T) | 0.53 | 9.38 | 6.58E-21 | *BCL11A* | 5 | | | | | ■ | ■ | ■ | ■ | | ■ |
| | 2q12.1 | Novel | rs62151809 (T:C) | 0.43 | 6.77 | 1.28E-11 | *POU3F3* | 3 | | | | | ■ | ■ | | ■ | | |
| | 2q24.3 | Replication | rs11890028 (T:G) | 0.72 | 5.63 | 1.73E-08 | *SCN1A* | 8 | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| | 2q32.2 | Replication | rs6721964 (A:G) | 0.66 | -6.18 | 6.54E-10 | *GLS* | 4 | | | | | | ■ | ■ | ■ | | ■ |
| | 3p22.3 | Novel | rs9861238 (A:G) | 0.41 | -6.42 | 1.33E-10 | *STAC* | 2 | | | | | ■ | | ■ | | | |
| | 3p21.31 | Novel | rs739431 (A:G) | 0.84 | 6.23 | 4.82E-10 | *CACNA2D2* | 6 | | | | ■ | | ■ | ■ | ■ | ■ | ■ |
| | 4p15.1 | Replication | rs1463849 (A:G) | 0.59 | -6.59 | 4.38E-11 | *PCDH7* | 3 | | | | ■ | ■ | | | | | |
| | 5q22.3 | Replication | rs4596374 (T:C) | 0.55 | -6.98 | 2.91E-12 | *KCNN2* | 6 | | | | ■ | ■ | ■ | ■ | ■ | | ■ |
| | 5q31.2 | Novel | rs2905552 (C:G) | 0.48 | -6.33 | 2.49E-10 | *SPOCK1* | 5 | | | | ■ | ■ | ■ | ■ | ■ | | |
| | 6q22.33 | Replication | rs13219424 (T:C) | 0.29 | -5.49 | 3.87E-08 | *PTPRK* | 3 | | | | | ■ | | ■ | ■ | | |
| | 7p14.1 | Novel | rs37276 (T:G) | 0.26 | -5.69 | 1.29E-08 | *SUGCT* | 2 | | ■ | | | ■ | | | | | |
| | 9q21.32 | Novel | rs2780103 (T:C) | 0.26 | -6.93 | 4.34E-12 | *RMI1* | 5 | ■ | ■ | ■ | ■ | | | | ■ | | |
| | 10q24.32 | Novel | rs11191156 (A:G) | 0.67 | -7.55 | 4.41E-14 | *KCNIP2* | 4 | | | | ■ | ■ | ■ | ■ | | | |
| | 12q13.13 | Novel | rs114131287 (A:T) | 0.02 | 5.83 | 5.46E-09 | *SCN8A* | 6 | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| | 16p13.3 | Novel | rs62014006 (T:G) | 0.05 | 5.88 | 4.22E-09 | *RBFOX1* | 5 | | | | ■ | ■ | ■ | ■ | ■ | | |
| | 17p13.1 | Novel | rs2585398 (A:C) | 0.53 | -6.37 | 1.84E-10 | *ARHGEF15* | 6 | ■ | ■ | ■ | | | | ■ | ■ | | |
| | 17q21.32 | Replication | rs16955463 (T:G) | 0.25 | -5.97 | 2.30E-09 | *CDK5RAP3* | 4 | | ■ | ■ | | | | | ■ | | |
| | 19p13.3 | Novel | rs75483641 (T:C) | 0.14 | -6.22 | 4.85E-10 | *AP3D1* | 5 | ■ | | ■ | | | | | ■ | | ■ |
| | 21q21.1 | Novel | rs1487946 (A:G) | 0.59 | 5.47 | 4.41E-08 | *TMPRSS15* | 1 | | | | | ■ | | | | | |
| | 21q22.1 | Replication | rs7277479 (A:G) | 0.36 | -6.82 | 8.94E-12 | *GRIK1* | 4 | | | | | | ■ | ■ | ■ | ■ | |
| | 22q13.32 | Novel | rs469999 (A:G) | 0.31 | -6.32 | 2.65E-10 | *FAM19A5* | 2 | | | | | ■ | ■ | | | | |
| CAE | 2p16.1 | Replication | rs12185644 (A:C) | 0.70 | -7.12 | 1.04E-12 | *BCL11A* | 5 | | | | | ■ | ■ | ■ | ■ | | ■ |
| JME | 4p12 | Replication | rs17537141 (T:C) | 0.851 | -5.47 | 4.62E-08 | *GABRA2* | 6 | | | | ■ | | ■ | ■ | ■ | ■ | |
| | 8q23.1 | Novel | rs3019359 (T:C) | 0.414 | -5.55 | 2.89E-08 | *RSPO2* | 3 | | | | | ■ | ■ | | | | |
| | | | | | | | *TMEM74* | 3 | | | | | | ■ | ■ | ■ | | |
| | 16p11.2 | Replication | rs1046276 (T:C) | 0.353 | 6.19 | 6.05E-10 | *STX1B* | 5 | | | | ■ | | ■ | ■ | ■ | | ■ |
| | | | | | | | *CACNA1I* | 5 | | | | | | ■ | ■ | ■ | ■ | ■ |

**Figure 2. Genome-wide significant loci and prioritized genes.** Genome-wise significant loci are annotated with details from the lead-SNP and prioritized genes. Genes were scored based on 10 criteria/methods, after which the gene with the highest score in the locus was selected as the prioritized gene. Total: number of satisfied criteria for gene prioritization. Missense: the locus contains a missense variant in the gene. TWAS: significant transcriptome-wide association with the gene. SMR: significant summary-based mendelian randomisation association with the gene. MAGMA: significant genome-wide gene based association. PoPS: gene prioritized by polygenic priority score. Brain exp: the gene is preferentially expressed in brain tissue. brain-coX: the gene is prioritized as co-expressed with established epilepsy genes. KO mouse: knockout of the gene causes a neurological phenotype in mouse models. Monogenic: the gene is a known monogenic epilepsy gene.

## The HLA and common epilepsies

We imputed HLA alleles and amino acid residues using COOKHLA and ran association across all epilepsy, focal, and GGE phenotypes (see Methods). No SNP, amino acid residue, or HLA allele reached the level of genome-wide significance. The most significant signal was with GGE, in which an aspartame amino acid residue in exon 2 position 31432494 had a p-value of $3.8e^{-07}$.

## SNP-based heritability

We calculated SNP-based heritability using LDAK, to determine the proportion of epilepsy risk attributable to common genetic variants. We observed liability scale SNP-based heritabilities of 17.7% (95%CI 15.5 - 19.9%) for all epilepsy, 16.0% (95%CI 14.0 - 18.0%) for focal epilepsy and 39.6% (34.3 - 44.6%) for generalized epilepsy. Heritabilities for GGE subtypes were notably higher for all individual GGE subtypes: ranging from 49.6% (14.0% - 85.3%) for GTCSA to 90.0% (63.3 - 116.6%) for JAE (**Supplementary table 5**).



Employing a univariate causal mixture model[24] (see Methods) we estimated that 2,849 causal SNPs (standard error: 199) underlie 90% of the SNP-based heritability of GGE. Power analysis demonstrated that the current genome-wide significant SNPs only explain 1.5% of the phenotypic variance, whereas a sample size of around 2.5 million subjects would be necessary to identify the causal SNP that explain 90% of generalized-epilepsy SNP-based heritability (**Supplementary figure 4**).

To further explore the heritability of the different epilepsy phenotypes, we used LDSC to perform genetic correlation analyses.[25] We found evidence for strong genetic correlation between all four GGE syndromes (**Supplementary figure 3**). We also observed a significant genetic correlation between the focal non-lesional and JME syndromes, which has been reported previously.[3] Here, with larger sample sizes, CAE also showed a significant genetic correlation with the focal non-lesional cohort.

## Tissue and cell-type enrichment

To further illuminate the biological aetiology of epilepsy, we used MAGMA and data from the gene-tissue expression consortium (GTEx) to assess

whether our GWAS-associated genes were enriched for expression in specific tissues and cell types (see Methods). We identified significant enrichment of associated genes expressed in brain and pituitary tissue (**Supplementary figure 5**). This is the first time the pituitary gland has been implicated in GGE, and might reflect a hormonal component to seizure susceptibility. Further sub-analyses showed that our results were enriched for genes expressed in almost all brain regions, including subcortical structures such as the hypothalamus, hippocampus and amygdala (**Supplementary figure 6**). We did not find enrichment for genes expressed at specific developmental stages in the brain (**Supplementary figure 7**).

Cell-type specificity analyses using various single-cell RNA-sequencing reference datasets (see Methods) revealed enrichment in excitatory as well as inhibitory neurons, but not in other brain cells like astrocytes, oligodendrocytes or microglia (**Supplementary figure 8**). Similarly, stratified LD-score regression using single-cell expression data (see Methods) did not reveal a difference between excitatory and inhibitory neurons (p=0.18).

## Gene-set analyses

MAGMA gene-set analyses showed significant associations between GGE and biological processes involving various functions in the synapse (**Supplementary data 7**). To further refine the synaptic signal, we performed a gene-set analysis using lists of expert curated gene-sets involving 18 different synaptic functions.[26] These analyses showed that GGE was associated with intracellular signal transduction (n=139 genes, $p=9.6e^{-5}$) and excitability in the synapse (n= 54 genes, p=0.0074). None of the other 16 synaptic functions showed any association (**Supplementary data 7**). Genes involved with excitability include the N-type calcium channel gene *CACNA2D2*, implicated at the novel GGE locus 3p21.31. N-type calcium channel blockers such as levetiracetam and lamotrigine are amongst the most widely used and effective anti-epileptic drugs for GGE as well as focal epilepsy.[27–29] Together, these results suggest that the genes associated with GGE are expressed in excitatory as well as inhibitory neurons in various brain regions, where they affect excitability and intracellular signal transduction in the synapse.

## Sex-specific analyses

There are known sex-related patterns in the epidemiology of epilepsy. Although females have a marginally lower incidence of epilepsy than males, GGE is known to occur more frequently in females.[30] In order to test whether this sex divergence has a genetic basis, we performed sex-specific GWAS for all, GGE and focal epilepsy (**Supplementary figures 9-11**). Analyses revealed one female-specific genome-wide significant signal at 10q24.32 (lead SNP:rs72845653), containing *KCNIP2*, implicated in our main GGE meta-analysis (lead SNP: rs11191156). However, the lead SNPs of these two signals are not in LD ($r^2$=0.05). Interestingly, the direction of effect of this signal is opposite in females and males. This sex difference is further corroborated by significant sex-heterogeneity (p=1.54e$^{-8}$) and gender-differentiation (p=5.6e$^{-9}$).[31] Sex-related differences in transcription levels in human heart have previously been reported for *KCNIP2*.[32] We did not find any sex-divergent signals for 'all' or focal epilepsy.

LDSC was used to assess the genetic correlation between male-only and female-only GWAS. The male and female GWAS of all epilepsy, focal and GGE were strongly genetically correlated (all Rg>0.9) and none of these correlations were significantly different from 1 (all p>0.05). These results suggest that, with the exception of the female-specific 10q24.32 signal, the overall genetic basis of common epilepsy appears largely similar between males and females.

## Genetic overlap between epilepsy and other phenotypes

In order to explore the genetic overlap of epilepsy with other diseases, we first cross-referenced the 26 genome wide epilepsy loci with other traits with significant associations (p<5x10^-8) for the same SNP, or SNPs in strong linkage disequilibrium with our lead SNPs (as detailed in **Figure 2**). This analysis revealed eighteen likely pleiotropic loci, with previous associations reported across a variety of traits, the most common being cognition, sleep, psychiatric, coronary and blood cell traits (**Supplementary figure 12**). The remaining eight loci appear to be specific to epilepsy (3p22.3, 4p12, 5q31.2, 7p14.1, 8q23.1, 9q21.13, 21q21.1, 21q22.1).

We then performed genetic correlation analyses between selected traits and all, GGE and focal epilepsy using LDSC[33]. The 17 selected traits had either or a combination of epilepsy as a common comorbidity or pleiotropic loci shared with epilepsy. Significant correlations, after correcting for multiple testing, were found with febrile seizures, stroke, ADHD, type 2 diabetes and intelligence, amongst others (**Figure 3**)

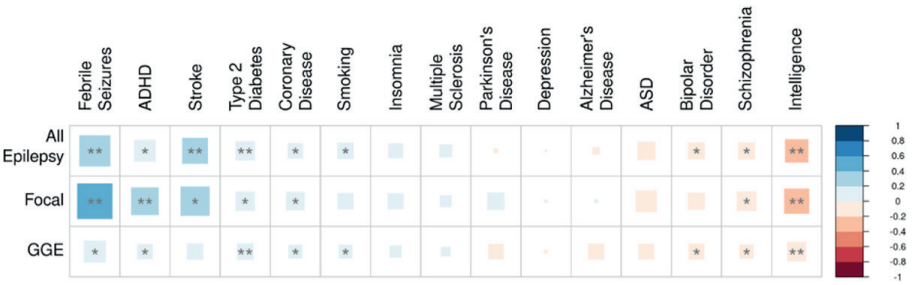



**Figure 3. Genetic correlations of epilepsy with other phenotypes**. The genetic correlation coefficient was calculated with LDSC and is denoted by color scale from −1 (red) to +1 (blue). * P < 0.05, ** P < 0.001 (Bonferroni corrected).

Genetic correlation analyses assess the aggregate of shared genetic variants associated with two phenotypes. However, genetic correlations can become close to zero when there is consistent mixed directionality of SNP effects between two phenotypes.[34] Traits such as ASD were not significantly correlated, despite monogenic pleiotropy with epilepsy genes supporting an overlap. To explore whether inverse directionality could explain the lack of genetic correlation between ASD and epilepsy we applied the MiXeR tool to generalized epilepsy, intelligence and autism spectrum disorder (ASD), to quantify polygenic overlap irrespective of genetic correlation (see Methods). Results showed that >99% of causal SNPs underlying generalized epilepsy are shared with intelligence, of which 58% have a discordant direction of effect (**Supplementary figure 13**). Furthermore, despite a lack of genetic correlation with ASD (Rg=−0.12, p=0.06, all epilepsy; Rg=−0.17, p=0.06 focal epilepsy; Rg=−0.09, p=0.09, GGE), we found that 95% of causal SNPs underlying generalized epilepsy are shared with ASD, but 59% have a discordant direction of effect. This is in line with the finding that epilepsy and ASD often co-occurs,[35] and monogenic forms of epilepsy and ASD can

have a shared genetic cause.[36] The discordant direction of effect is in line with evidence of monogenic causes of ASD and epilepsy due to pathogenic variants in *SCN2A*. Functional studies have shown that ASD without seizures can be caused by loss-of-function variants in *SCN2A* [37], whereas epilepsy can be caused by gain-of-function variants in *SCN2A*.[38,39] Indeed, ASD variants in *SCN2A* seem protective against neuronal hyperexcitability.[39]

## Leveraging GWAS for drug repurposing

In order to test the potential of our meta analysis to inform drug repurposing, we utilized a method that predicts the relative efficacy of a drug for epilepsy, based upon that drugs' predicted ability to modulate epilepsy-related changes in the function and abundance of proteins, as inferred from the GWAS summary statistics (see Methods).[40] We validated the drug predictions by determining if they are concordant with findings from clinical experience and trials. In our predictions for all epilepsy, current anti-seizure drugs were ranked higher than expected by chance ($p < 1\times10^{-6}$), and higher than drugs used to treat any other human disease. For GGE, broad-spectrum antiseizure drugs were predicted to be more effective than narrow-spectrum antiseizure drugs ($p < 1\times10^{-6}$), consistent with clinical experience.[41] Furthermore, the predicted order of efficacy for GGE of *individual* antiseizure drugs' matched their observed order in the largest head-to-head randomized controlled clinical trials for generalized epilepsy, [29,42] an observation is unlikely to occur by chance ($p < 1\times10^{-6}$).

Using this approach, we highlight the top 20 drugs that are licensed for conditions other than epilepsy, but are predicted to be efficacious for generalised epilepsy, *and* additionally have published evidence of anti-seizure efficacy from multiple published studies and multiple animal models (**Supplementary table 6**). The fill list of all predictions can be found in **Supplementary data 8**.

## Biobank results

We performed GWAS using data from several large-scale population biobanks (total cases n=21,734, total controls n=1,023,989, phenotyped using ICD codes, see Methods). Although the biobank-specific GWAS did not identify any genome-wide significant loci for GGE or 'all epilepsy',

one significant locus at 2q22.1 emerged for focal epilepsy (**Supplementary figure 14**).

Meta-analysing the biobank summary statistics with those from the primary epilepsy GWAS identified seven significant loci for the 'all epilepsy' phenotype. Six of these signals were previously identified in the primary 'all epilepsy' (n=4) or the 'GGE' GWAS (n=2). One locus (2q12.1) was novel. The combined biobank meta-analysis for GGE identified five novel loci, but four loci from our primary GWAS fell below significance (**Supplementary figure 15**). The combined focal epilepsy meta-analysis showed no significant associations. LDSC between the biobank-only and the primary GWAS results showed genetic correlations ranging between 0.31 and 0.74 (**Supplementary table 7**).

# Discussion

In this study, we leveraged a substantial increase in sample size to uncover 26 common epilepsy risk loci, of which 16 have not been reported previously. Using a combination of ten post GWAS analysis methods, we were able to pinpoint 29 genes that most likely underlie these signals of association. These signals showed enrichment throughout the brain and indicate an important role for synapse biology in excitatory as well as inhibitory neurons. Drug prioritization from the genetic data highlighted licensed AEDs, ranked the AEDs broadly in line with clinical experience and pointed to drugs for potential repurposing. These findings further our understanding of the pathophysiology of common epilepsies and provide new leads for therapeutics.

The 26 associated loci included some notable novel epilepsy genes. These include the calcium channel gene *CACNA2D2*, an established epileptic encephalopathy gene[43] and directly targeted by ten currently licenced drugs, including two anti-epileptic drugs (gabapentin and pregabalin) as well as the Parkinson's disease drug safinamide and the nonsteroidal anti-inflammatory drug celecoxib. Both safinamide and celecoxib have evidence of anti-seizure activity.[44,45] *SCN8A*, which encodes a voltage-gated sodium channel, is an established epileptic encephalopathy gene and is associated

with common epilepsies for the first time here. *SCN8A* is targeted by drugs including safinamide and quinidine. *RYR2* encodes a ryanodine receptor, is an established cardiac disorder gene, has recently been implicated in epilepsy [46,47] and is targeted by caffeine as well simvastatin, atorvastatin and carvedilol. The acetylcholine receptor gene *CHRM3* has not been previously associated with epilepsy and is targeted by drugs including solifenacin, used to treat urinary incontinence.

We found that GGE in particular has a relatively strong contribution from common genetic variation. When analyzing individual GGE syndromes, we found that up to 90% of liability is attributable to common variants, which is higher than any of 778 other traits studied in a large GWAS atlas.[48] For the collective GGE phenotype, heritability estimates decrease to 40%, which is still higher than 773/778 other traits.[48] This decrease could be explained by increased heterogeneity, from combining syndromes with pleiotropic as well as syndrome-specific risk loci. Although statistical power drastically decreased when assessing specific GGE syndromes, three loci appeared specific to JME. These findings highlight the unique genetic architecture of the subtypes of common epilepsies, which are characterized by a high degree of both shared, and syndrome-specific, genetic risk.



In contrast to GGE, for focal epilepsies we found only a minor contribution of common variants, with no variant reaching genome-wide significance. It would seem that focal epilepsies, as a group, are far more heterogeneous than GGE. Our attempt to mitigate this heterogeneity by performing subtype analysis, contrasted with the results from GGE, suggesting different genetic architectures, consistent with the experience from studies of rare genetic variation and PRS.[4,5]

This work highlights the challenges of working with epilepsy cohorts ascertained through large biobanking initiatives. Accurate classification of epilepsy requires a combination of imaging, electrophysiology and clinical features. These details were not available from the biobanks we worked with. Rather, phenotypes were generally limited to ICD codes, which are prone to misclassification.[49] Population biobanks are also probably ascertaining for milder epilepsies that are responsive to treatment, contrasting with the

enrichment for refractory epilepsies at tertiary referral centres. Moreover, a large proportion of people with epilepsy have a less-heritable etiology, such as epilepsy due to stroke and tumors and head trauma, and based on available phenotype, such cases could not be identified and excluded from the biobanks studied here. As a result, the inclusion of the biobank data appeared to introduce more heterogeneity. This contrasts with genetic mapping of other polygenic diseases like type 2 diabetes and migraine, which are relatively easy and reliable to diagnose, resulting in a great increase in GWAS loci when including data from the same biobanks as included in our study.[83,84]

We found enrichment of GGE variants in brain-expressed genes, involving excitatory and inhibitory neurons, but not any other brain cell. This contrasts with other neurological diseases. For example, microglia are primarily involved in Alzheimer's disease[50] and multiple sclerosis,[51] whereas migraine does not appear to have brain cell specificity.[52] We further refine this signal by showing an involvement of synapse biology, primarily intracellular signal transduction and synapse excitability. These findings suggest an important role of synaptic processes in excitatory and inhibitory neurons throughout the brain, which could be a potential therapeutic target. Indeed synaptic vesicle transport is a known target of the AEDs levetiracetam and brivaracetam.[85]

We confirmed that the identified genes in our GWAS overlapped with monogenic epilepsy genes and known targets of current anti-epileptic drugs.[3] We extend this observation by providing a list of other drugs that directly target the genes prioritized in our GWAS. Moreover, using a systems-based approach[40] we highlight drugs that are predicted to be efficacious when repurposed for epilepsy, based on their ability to perturb function and abundance in gene expression. Insights from GWAS of epilepsy has the potential to accelerate clinical trials, via the identification of promising drug repurposing candidates for clinical trials.[53] We anticipate that follow-up studies of the highlighted drugs in this study could show clinical efficacy in epilepsy treatment.

In summary, these new data reveal markedly different genetic architectures between the milder and more common focal and generalized epilepsies,

provide novel biological insights to disease aetiology and highlight drugs with predicted efficacy when repurposed for epilepsy treatment.

# Methods

## Ethics statement

Local institutional review boards approved study protocols at each contributing site. All study participants provided written, informed consent for use of their data in genetic studies of epilepsy. For minors, written informed consent was obtained from their parents or legal guardian.

## Sample and phenotype descriptions



This meta-analysis combines previously published datasets with novel genotyped cohorts. Descriptions of the 24 cohorts included in our previous analysis can be found in the Supplementary table 6 of that publication.[3] Here we included 5 novel cohorts (**see Supplementary table 8**), comprising 14,732 epilepsy cases and 22,861 controls, resulting in a total sample size of 27,559 cases and 42,436 controls. Classification of epilepsy was performed as described previously.[3] In brief, we assigned people with epilepsy into focal epilepsy, genetic generalised epilepsy (GGE) or unclassified epilepsy. 'All epilepsy' was the combination of GGE, focal and unclassified epilepsy. Where possible, we used EEG, MRI and clinical history to further refine the subphenotypes: juvenile myoclonic epilepsy (JME), childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), generalised tonic-clonic seizures alone (GTCS), non-lesional focal epilepsy, focal epilepsy with hippocampal sclerosis (HS) and focal epilepsy with lesion other than HS.

## Genotyping, quality control and imputation

Subjects were genotyped on single nucleotide polymorphism (SNP) arrays, see **Supplementary table 7** for an overview of genotyping in novel cohorts. Quality control (QC) was performed separately for each cohort. Prior to imputation, data from the Janssen, Austrian, Swiss, Norwegian, and BPCCC cohorts were cross-referenced to the HRC panel to ensure SNPs matched

in terms of strand, position, and ref/alt allele assignment. Additionally, SNPs were removed if they were absent in the HRC panel, if they had a >20% allele frequency difference with the HRC panel, or if any AT/GC SNPs had MAFs>40%, using tools available from https://www.well.ox.ac.uk/~wrayner/tools/ . Data were then imputed using the the Wellcome Sanger Institutes' imputation server (https://imputation.sanger.ac.uk/), using EAGLE v2.4.1 (Loh *et al.,* 2016) for phasing, and the Positional Burrows Wheeler Transform algorithm (Rubinacci *et al.,* 2020) for imputation. The Haplotype Reference Consortium (HRC) reference panel r1.1, was used as a reference for imputation (Haplotype Reference Consortium, 2016). Post-imputation, SNPs with an INFO score of <0.9 were removed. The high-INFO SNPs were then converted back to PLINK format and once-again QC'd for genotype coverage (>0.98), minor allele frequencies (>5%) and Hardy-Weinberg Equilibrium violations (p>10e-5), following previously described methodologies[3].

QC for the Epi25 cohort was performed using a similar in-house pipeline. Samples were split by ethnicity based on principal component analysis. Pre-imputation QC included filtering of SNPs with call rate (<98%), differential missing rate, duplicated and monomorphic SNPS, SNPs with batch association (p<1e-4), violation of Hardy-Weinberg Equilibrium (p<1e-10). Sample filtering included removal of outliers (>4SD from mean) of heterozygous/homozygous ratio, removal of one of each pair of related samples (proportion identity-by-descent >0.2) and removal of samples with ambiguous or non-matching genetically imputed sex. Furthermore, duplicates between the Epi25 cohort and the previously published genome-wide mega-analysis were identified based on genotype, after which these subjects were removed from the Epi25 cohort. Genotypes were imputed on the Michigan imputation server, using the Haplotype Reference Consortium v1.1 (n=32470) reference panel for subjects of European and Asian ancestry, and the 1000 Genomes Phase 3 v5 (n=2504) reference panel for subjects of African ancestry. Default imputation parameters and pre-imputation checks were used. Imputed dosages were used for subsequent analyses, filtering on imputation INFO>0.3 and minor-allele frequency >0.01.

## Genome-wide association analyses

GWAS of the Janssen Pharmaceuticals, Swiss GenEpa, Norwegian GenEpa and Austrian GenEpa cohorts was performed as described previously.[3] GWAS of the Epi25 cohort was performed with a generalized mixed model using SAIGE v0.38.[54] SAIGE was performed in two steps: (1) fitting the null logistic mixed model to estimate the variance component and other model parameters; (2) testing for the association between each genetic variant and phenotypes by applying SPA to the score test statistics. For step 1, SNPs were filtered on call rate >0.98 and MAF >5%, and SNPs were pruned to obtain approximate independent markers (window size of 100kb and R2>0.3), while including sex and the top 10 principal components as covariates. Next, we performed fixed-effects meta-analyses with METAL,[55] for each of the main phenotypes (all, GGE, and focal epilepsy), as well as the subphenotypes, weighted by effective samples sizes ($N_{eff} = 4/(1/N_{cases} + 1/N_{controls})$) to account for case-control imbalance. We performed trans-ethnic and European-only meta-analyses for the main phenotypes, and restricted the subphenotype analyses to Europeans only, due to limited sample size in other ethnicities. We included all SNPs (~4.9 million) that were present in at least the previous mega-analysis and the Epi25 dataset, which together account for 88% of the total sample size. We calculated genomic inflation factors ($\lambda$), mean $\chi 2$ and LD score regression intercepts to assess potential inflation of the test statistic. Since $\lambda$ is known to scale with sample size, we also calculated $\lambda 1000$, which is $\lambda$ corrected for an equivalent sample size of 1000 cases and 1000 controls.[56] We limited these analyses to subjects of European ancestry, since these LD-structure depends on ethnicity and Europeans constituted 92% of cases.

## Biobank GWAS

We had access to summary statistics of epilepsy GWAS from four population biobanks; UK Biobank (Sudlow *et al.,* 2015), Biobank Japan (Nagai *et al.,* 2017), Finngen release R6 (Finland; https://finngen.gitbook.io/), and DECODE genetics (Iceland; (Gudbjartsson *et al.,* 2015)). The biobank Japan, Finngen and DECODE genetic cases were assigned into either 'focal' or 'generalised' epilepsy, whereas the UK Biobank samples were not subdivided based on

seizure localisation, as the relevant clinical details were unavailable to facilitate an accurate sub division.

## Control data were population matched samples with no history of epilepsy.

GWAS fixed-effects meta-analysis were conducted using METAL (Willer *et al.,* 2010). To account for case-control imbalance the effective sample size for each cohort was calculated as $N_{eff} = 4/(1/N_{cases} + 1/N_{controls})$. GWAS Manhattan plots and QQ plots were generated using the qqman R package (Turner, 2014). Genome-wide significant loci were mapped onto genes using the FUMA web platform (Watanabe *et al.,* 2017).

We performed three meta-analyses. As a primary analysis, we meta-analysed all non-biobank samples, then we meta-analysed only biobank samples and finally performed a combined meta-analysis of biobank and non-biobank samples.

## Pleiotropy analysis

For a dedicated pleiotropy detection analysis, we used the ASSET method.[86] ASSET is a meta-analysis-based pleiotropy detection approach that identifies common or shared genetic effects between two or more related but distinct traits. It identifies variants that are associated with a subset of traits (or even all traits) as well as their direction of effect, thereby extending the classical meta-analysis by considering subsets of phenotypes. ASSET estimates the evidence for pleiotropy by $Z_{meta-max}$ which denotes the maximum effect of a SNP from all the associated subsets of trait(s). Corresponding P-values are obtained through the discrete local maxima (DLM) method. We used ASSET with a genome-wide significance level of $\alpha=5\times10^{-8}$.

We applied ASSET to the subset of European samples, comprising 6952 (3244+3708) GGE cases and 14,939 (5344+9095) FE cases from the EPI25 and our Consortium as well as 42,434 partially overlapping controls from both consortia. To this end, we used effect sizes, standard errors and the effective sample sizes estimated from the main meta-analysis. Note that ASSET accounts for sample overlap in the analysis.

## HLA association

HLA types and amino acid residues were imputed using the COOKHLA software,[58] with the 1000 genomes phase 3 used as a reference panel.[59] Samples were grouped by genetic ancestry for imputation.

Following imputation, association analysis was conducted using the HLA Analysis Toolkit (HATK) (Choi *et al.,* 2021).[60] Three phenotypes were analysed: 'all epilepsy', 'focal epilepsy', and 'GGE'. Samples from the ILAE and Epi25 datasets were analysed separately, and the association results were meta-analysed across datasets using PLINK.[61]

## Functional annotation

We annotated all genome-wide significant SNPs and tagged SNPs within the loci. ANNOVAR was used to annotate the location and function of each SNP,[62] the CADD score was used as a measure of predicted deleteriousness,[63] and chromatin states were annotated using epigenetic data from the ENCODE and NIH Roadmap Epigenomics Mapping Consortium.[12,64] We used FUMA to define independently significant SNPs within loci; i.e. SNPs that were genome-wide significant but not in LD (R2<0.2 in Europeans) with the lead SNP in the locus.

## Gene mapping

We used FUMA[16] to map genome-wide significant loci to specific genes, using the same parameters as published previously.[3] We defined genome-wide significant loci as the region encompassing all SNPs with P<1e-4 that were in LD (R2>0.2) with the lead SNP (i.e. the SNP with the strongest association within the region). We used a combination of positional mapping (within 250kb from locus), eQTL mapping (SNPs with FDR corrected eQTL P<0.05 in blood or brain tissue) and 3D Chromatin Interaction Mapping (FDR p <1e-6 in brain tissue).

## Genome-wide gene based association study and gene-set analyses

MAGMA v1.6 was used to perform a genome-wide gene based association study (GWGAS) and gene-set analyses.[17] GWGAS was performed using

default settings of MAGMA, as implemented in FUMA, which calculates an association P-value based on all the associations of all SNPs within each gene in the GWAS. Based on these GWGAS results, we performed competitive gene-set analyses with default MAGMA settings, using 15483 default gene sets and GO-terms from MsigDB. In addition, we specifically assessed 18 curated gene-sets involving different synaptic functions.[26]

## Transcriptome wide association study

Transcriptome wide association studies (TWAS) were performed with FUSION v3, with default settings.[18] We imputed gene expression based on our European-only GWAS (since the method relies on LD reference data) eQTL data from the PsychENCODE consortium, which includes dorsolateral prefrontal cortex tissue from 1695 human subjects.[19]

## Summary-data-based Mendelian Randomization

We used Summary-data-based Mendelian Randomization (SMR) v1.03 as an additional method to assess the association between epilepsy and expression of specific genes.[20] SMR tests whether the effect size of a SNP on epilepsy is mediated by expression of specific genes. We performed SMR analyses with default settings, using the MetaBrain expression data as a reference; a new eQTL dataset including 2970 human brain samples.[65]

## Sex-specific analyses

We performed sex-specific and sex-divergence GWAS to assess potential signals that are specific to males or females. First, we performed the same GWAS as described above for all epilepsy (13889 female cases and 19676 female controls; 12259 male cases and 18645 male controls) and GGE (3946 female cases and 19676 female controls; 2603 male cases and 18645 male controls) seperately for subjects of either sex, after which we performed fixed-effects meta-analyses with METAL to merge the different cohorts. Next, we performed meta-analyses between the male and female GWAS with GWAMA[66] to assess heterogeneity of effect sizes between sexes and to calculate gender-differentiated associations.[31]

## Gene prioritization

We used a combination of 10 different methods to find the most likely biological candidate gene within each genome-wide significant locus. For each gene in each locus we assessed the following criteria:

- Missense: we assessed whether the SNPs tagged in the genome-wide significant locus contained an exonic missense variant in the gene, as annotated by ANNOVAR.
- TWAS: we assessed whether imputed gene expression was significantly associated with the epilepsy phenotype, based on the FUSION TWAS as described above, Bonferroni corrected for each mapped gene with expression information.
- SMR: we assessed whether the gene had a significant SMR association with the epilepsy phenotype, based on the SMR analyses as described above, Bonferroni corrected for each mapped gene with expression information.
- MAGMA: we assessed whether the gene was significantly associated with the epilepsy phenotype through a GWGAS analysis, Bonferroni corrected for each mapped gene.
- PoPS: we calculated the Polygenic Priority Score (PoPS)[67]; a novel method that combines GWAS summary statistics with biological pathways, gene expression, and protein-protein interaction data, to pinpoint the most likely causal genes. We scored the gene with the highest PoPS score within each locus.
- Brain expression: we calculated mean expression of all brain and non-brain tissues based on data from the Genetype-Tissue Expression (GTEx) project v8[68] and assessed if the average brain tissue expression was higher than the average expression in non-brain tissues.
- brain-coX: we assessed whether genes were prioritized as co-expressed with established epilepsy genes in more than a third of brain tissue resources utilized, using the tool brain-coX (**Supplementary figure 16**).[69]
- Target of AED: we assessed whether the gene is a known target of an anti-epileptic drug, as assessed in the drug-gene interaction database (www.DGidb.com; accessed on 26-11-2021) and a list of drug targets from a recent publication.[70]

- Knockout mouse: we assessed whether a knockout of the gene in a mouse model results in a nervous system (phenotype ID: MP:0003631) or a neurological/behaviour phenotype (MP:0005386) in the Mouse Genome Informatics database (http://www.informatics.jax.org; accessed on 26-11-2021).
- Monogenic epilepsy gene: we evaluated whether the gene is listed as a monogenic epilepsy gene, in a curated list maintained by the Epilepsy Research Centre at the University of Melbourne.

Similar to previous studies,[3,71] we scored all genes based on the number of criteria being met (range 0-10; all criteria had an equal weight). The gene with the highest score was chosen as the most likely implicated gene. We implicated two genes if both had an identical score.

## Long distance expression regulation of BCL11A

Most eQTL databases like PsychENCODE and MetaBrain restrict eQTL analyses to 1MB distance between genes and SNPs. To specifically assess the hypothesis of long-distance regulation of BCL11A by the lead SNPs in the 2p16.1 epilepsy locus, we manually interrogated the MetaBrain database[65] without distance restraints. Next, we calculated the association between the 3 lead SNPs in the locus (rs11688767, rs77876353, rs13416557) with BCL11A expression.

## Heritability analyses

We performed LDAK analyses to calculate SNP-based heritability, using default settings with pre-calculated LD weights from 2000 European (white British) reference samples under the BLD-LDAK SumHer model, as recommended for human traits.[72] We performed these analyses for the main epilepsy phenotypes and subphenotypes, based on our European-only GWAS. SNP based heritabilities were converted to liability scale heritability estimates, using the formula: h2L=h2o*K2(1–K)2/p(1–p)*Z2, where K is the disease prevalence, p is the proportion of cases in the sample, and Z is the standard normal density at the liability threshold. As recently suggested to decrease downward bias, we performed these calculations based on the

effective sample sizes (see calculation above), after which p=0.5 can be assumed.[73] We used the same population prevalences as our previous study.[3]

We used a causal mixture model (MiXeR) to estimate the total amount of causally associated variants (i.e. variants with nonzero additive genetic effect) underlying epilepsy risk.[34] MiXeR utilizes a likelihood-based framework to estimate the amount of causal SNPs underlying a trait, without the need to pinpoint which specific SNPs are involved. Furthermore, MiXeR allows for power calculations to assess the required sample size to explain a certain proportion SNP-based heritability by genome-wide significant SNPs.

## Enrichment analyses

We used MAGMA, as implemented in FUMA, to perform tissue and cell-type enrichment. First, we assessed whether our GGE GWAS was enriched for specific tissues from the GTEx database. Similarly, we assessed enrichment of genes expressed in the brain at 11 general developmental stages, using data from the BrainSpan consortium. Next, we assessed whether GGE was associated with specific cell types, by cross-referencing two single-cell RNA sequencing databases of human developmental and adult brain samples. The PsychENCODE database contains RNA sequencing data from 4249 human brain cells from developmental stages and 27412 human adult brain cells.[74] The Zhong dataset (GSE104276) contains RNA sequencing data from 2309 human brain cells at different stages in development.[75] We performed FDR correction across datasets to assess which cell types were significantly associated with GGE. As sensitivity analysis, we performed stratified LDSC with default settings using the cell-specific gene expression weights from the PsychENCODE consortium to compare GABAergic with glutamatergic neuron enrichment.[76]

## Genetic overlap with other diseases

Using the FUMA web application, we searched the GWAS Catalog for previously reported associations with $P < 5 * 10^{-8}$ for SNPs at all 26 genome-wide significant loci.

Genetic correlations between all, focal epilepsy and GGE and other traits were computed with LDSC, using default settings. Traits highlighted by the GWAS catalog analysis and/or those with established epilepsy comorbidity were prioritised and pursued provided recent summary statistics were available for public download (**Supplementary table 9**).

We used a recently described bivariate causal mixture model to quantify polygenic overlap between GGE with intelligence and autism spectrum disorder (ASD). Publicly available summary statistics from intelligence (n=269867) and ASD GWAS (n=46350) were downloaded,[77,78] after which bivariate MiXeR was run with default settings.

## Drug-repurposing analyses

We hypothesised that the results of this GWAS can be used to predict drugs that have antiseizure efficacy. We utilised a recently developed method that uses the GWAS for a disease to predict the relative efficacy of drugs for the disease.[40] This method predicts the relative efficacy of drugs for a disease based upon drugs' inferred ability to modulate changes in the function and abundance of proteins caused by common genetic variations associated with the disease. We applied this method to the all epilepsy and GGE GWAS results, using (1) imputed gene expression data from the FUSION analyses, as described above, and (2) gene-based p-values from GWGAS analyses with MAGMA (see above), with default settings. We validated the drug predictions by determining if they are concordant with findings from clinical experience and trials. We determined if our predictions correctly identify (area under receiver operating characteristic curve) and prioritize (median rank) known clinically-effective antiseizure drugs, as previously described.[40] We determined the statistical significance of drug identification and prioritization results by comparing the results to those from a null distribution generated by performing $10^6$ random permutations of the scores assigned to drugs. Finally, we produced a list of the top 20 drugs predicted for generalized epilepsy, which are currently licensed for conditions other than epilepsy, but have published evidence of antiseizure efficacy from multiple studies and in multiple animal models.

### Biobank GWAS

We obtained summary statistics of epilepsy GWAS from four population biobanks; UK Bioban,[79] Biobank Japan,[80] Finngen release R6,[81] and DECODE genetics[82] (Iceland). The biobank Japan, Finngen and DECODE genetic cases were assigned into either 'focal' or 'generalised' epilepsy, whereas the UK Biobank samples were not subdivided based on seizure localisation, as the relevant clinical details were unavailable to facilitate an accurate sub division (see **Supplementary table 10** for sample sizes per biobank). Control data were population matched samples with no history of epilepsy.

Fixed-effects meta-analysis were conducted using METAL[55], weighted by effective sample size ($N_{eff} = 4/(1/N_{cases} + 1/N_{controls})$) to account for case-control imbalance.



We first performed meta-analyses on biobank-only samples, after which we used METAL to perform a meta-analyses to combine our main GWAS with the biobanks.

## Supplementary materials

Supplementary materials can be found at: https://tinyurl.com/4j3srm5j

## Authors

The International League Against Epilepsy Consortium on Complex Epilepsies*

*three lead analysts listed first followed by all members in alphabetical order

Remi Stevelink, Ciarán Campbell, Siwei Chen, Oluyomi M Adesoji, Zaid Afawi, Elisabetta Amadori, Alison Anderson, Joseph Anderson, Danielle M Andrade, Grazia Annesi, Andreja Avbersek, Melanie Bahlo, Mark D Baker, Ganna Balagura, Simona Balestrini, Carmen Barba, Karen Barboza, Fabrice Bartolomei, Thomas Bast, Larry Baum, Tobias Baumgartner, Betül Baykan, Nerses Bebek, Albert J Becker, Felicitas Becker, Caitlin A Bennett, Bianca Berghuis, Samuel F Berkovic, Ahmad Beydoun, Claudia Bianchini,

Francesca Bisulli, Ilan Blatt, Ingo Borggraefe, Christian Bosselmann, Vera Braatz, Jonathan P Bradfield, Knut Brockmann, Lawrence C Brody, Russell J Buono, Robyn M Busch, Hande Caglayan, Ellen Campbell, Laura Canafoglia, Christina Canavati, Gregory D Cascino, Barbara Castellotti, Claudia B Catarino, Gianpiero L Cavalleri, Felecia Cerrato, Francine Chassoux, Stacey S Cherny, Krishna Chinthapalli, I-Jun Chou, Seo-Kyung Chung, Claire Churchhouse, Peggy O Clark, Andrew J Cole, Alastair Compston, Antonietta Coppola, Mahgenn Cosico, Patrick Cossette, John J Craig, Caroline Cusick, Mark J Daly, Lea K Davis, Gerrit-Jan de Haan, Norman Delanty, Chantal Depondt, Philippe Derambure, Orrin Devinsky, Lidia Di Vito, Dennis J Dlugos, Viola Doccini, Colin P Doherty, Hany El-Naggar, Christian E Elger, Colin A Ellis, Johan G Eriksson, Annika Faucon, Yen-Chen A Feng, Lisa Ferguson, Thomas N Ferraro, Lorenzo Ferri, Martha Feucht, Mark Fitzgerald, Beata Fonferko-Shadrach, Francesco Fortunato, Silvana Franceschetti, Andre Franke, Jacqueline A French, Elena Freri, Monica Gagliardi, Antonio Gambardella, Eric B Geller, Tania Giangregorio, Leif Gjerstad, Tracy Glauser, Ethan Goldberg, Alicia Goldman, Tiziana Granata, David A Greenberg, Renzo Guerrini, Namrata Gupta, Hakon Hakonarson, Kerstin Hallmann, Manu Hegde, Erin L Heinzen, Ingo Helbig, Christian Hengsbach, Henrike O Heyne, Shinichi Hirose, Edouard Hirsch, Helle Hjalgrim, Daniel P Howrigan, Po-Cheng Hung, Michele Iacomino, Lukas L Imbach, Yushi Inoue, Atsushi Ishii, Jennifer Jamnadas-Khoda, Lara Jehi, Michael R Johnson, Reetta Kälviäinen, Yoichiro Kamatani, Moien Kanaan, Masahiro Kanai, Anne-Mari Kantanen, Bülent Kara, Symon M Kariuki, Dalia Kasperavičiūte, Dorothee Kasteleijn-Nolst Trenite, Mitsuhiro Kato, Josua Kegele, Yeşim Kesim, Nathalie Khoueiry-Zgheib, Chontelle King, Heidi E Kirsch, Karl M Klein, Gerhard Kluger, Susanne Knake, Robert C Knowlton, Bobby P C Koeleman, Amos D Korczyn, Andreas Koupparis, Ioanna Kousiappa, Roland Krause, Martin Krenn, Heinz Krestel, Ilona Krey, Wolfram S Kunz, Mitja I Kurki, Gerhard Kurlemann, Ruben Kuzniecky, Patrick Kwan, Angelo Labate, Austin Lacey, Dennis Lal, Zied Landoulsi, Yu-Lung Lau, Stephen Lauxmann, Stephanie L Leech, Anna-Elina Lehesjoki, Johannes R Lemke, Holger Lerche, Gaetan Lesca, Costin Leu, Naomi Lewin, David Lewis-Smith, Qingqin S Li, Laura Licchetta, Kuang-Lin Lin, Dick Lindhout, Tarja Linnankivi, Iscia Lopes-Cendes, Daniel H Lowenstein, Colin H T Lui, Francesca Madia, Sigurdur

Magnusson, Anthony G Marson, Patrick May, Christopher M McGraw, Davide Mei, James L Mills, Raffaella Minardi, Nasir Mirza, Rikke S Møller, Anne M Molloy, Martino Montomoli, Barbara Mostacci, Lorenzo Muccioli, Hiltrud Muhle, Karen Müller-Schlüter, Imad M Najm, Wassim Nasreddine, Benjamin M Neale, Bernd Neubauer, Charles RJC Newton, Markus M Nöthen, Michael Nothnagel, Peter Nürnberg, Terence J O'Brien, Yukinori Okada, Elías Ólafsson, Karen L Oliver, Çiğdem Özkara, Aarno Palotie, Faith Pangilinan, Savvas S Papacostas, Elena Parrini, Manuela Pendziwiat, Slavé Petrovski, William O Pickrell, Rebecca Pinsky, Tommaso Pippucci, Annapurna Poduri, Federica Pondrelli, Rob H W Powell, Michael Privitera, Annika Rademacher, Rodney Radtke, Francesca Ragona, Sarah Rau, Mark I Rees, Brigid M Regan, Philipp S Reif, Sylvain Rhelms, Antonella Riva, Felix Rosenow, Philippe Ryvlin, Anni Saarela, Lynette G Sadleir, Josemir W Sander, Thomas Sander, Marcello Scala, Theresa Scattergood, Steven C Schachter, Christoph J Schankin, Ingrid E Scheffer, Bettina Schmitz, Susanne Schoch, Susanne Schubert-Bast, Andreas Schulze-Bonhage, Paolo Scudieri, Beth R Sheidley, Jerry J Shih, Graeme J Sills, Sanjay M Sisodiya, Michael C Smith, Philip E Smith, Anja C M Sonsma, Doug Speed, Michael R Sperling, Hreinn Stefansson, Kári Stefansson, Bernhard J Steinhoff, Ulrich Stephani, William C Stewart, Carlotta Stipa, Pasquale Striano, Hans Stroink, Adam Strzelczyk, Rainer Surges, Toshimitsu Suzuki, K Meng Tan, George A Tanteles, Erik Taubøll, Liu Lin Thio, Rhys H Thomas, Oskari Timonen, Paolo Tinuper, Marian Todaro, Pınar Topaloğlu, Rossana Tozzi, Meng-Han Tsai, Birute Tumiene, Dilsad Turkdogan, Unnur Unnsteinsdóttir, Algirdas Utkus, Priya Vaidiswaran, Luc Valton, Andreas van Baalen, Annalisa Vetro, Eileen P G Vining, Frank Visscher, Sophie von Brauchitsch, Randi von Wrede, Ryan G Wagner, Yvonne G Weber, Sarah Weckhuysen, Judith Weisenberg, Michael Weller, Peter Widdess-Walsh, Markus Wolff, Stefan Wolking, David Wu, Kazuhiro Yamakawa, Wanling Yang, Zuhal Yapıcı, Emrah Yücesan, Sara Zagaglia, Felix Zahnert, Federico Zara, Wei Zhou, Fritz Zimprich, Gábor Zsurka, Quratulain Zulfiqar Ali

# Author contributions

## Data analysis:

*Analytical design, imputation:* O.M.Adesoji, M.Bahlo, C.Campbell (lead analyst), G.L.Cavalleri, S.Chen (lead analyst), Y-C.A.Feng, B.P.C.Koeleman, R.Krause (data management), D.Lal, C.Leu, N.Mirza, M.Nothnagel, K.L.Oliver, R.Stevelink (lead analyst).

*Data generation and quality control and management:* L.Baum, J.P.Bradfield, R.J.Buono, G.L.Cavalleri., F.Cerrato, S.S.Cherny, C.Churchhouse, C.Cusick, Y-C.A.Feng, N.Gupta, H.Hakonarson, E.L.Heinzen, I.Helbig, D.P.Howrigan, D.Kasperaviciute, B.P.C.Koeleman, R.Krause., D.Lal, Z.Landoulsi, C.Leu, I.Lopes-Cendes., P.May, N.Mirza, B.M.Neale, P.-W.Ng, P.Nürnberg, Sl.Petrovski, T.Sander, D.Speed, R.Stevelink, Fe.Zara, W.Zhou.

*External data resources and analysis:* UK BioBank: C.Campbell, D.Lewis-Smith, R.H.Thomas.

BioBank Japan: Y.Kamatani, M.Kanai, M.Kato, Y.Okada.

FinnGenn: M.J.Daly, H.O.Heyne, R.Kälviäinen, M.I.Kurki, A.Palotie.

deCODE genetics: S.Magnusson, E.Ólafsson, H.Stefansson, K.Stefansson, U.Unnsteinsdóttir.

*Analysis coordination:* G.L.Cavalleri (Co-Chair), B.P.C.Koeleman (Co-Chair)

**Writing committee:** O.M.Adesoji, M.Bahlo, S.F.Berkovic, C.Campbell, G.L.Cavalleri, S.Chen, B.P.C.Koeleman, K.L.Oliver, R.Stevelink (wrote first draft).

**Strategy committee:** L.Baum, S.F.Berkovic (Chair), R.J.Buono, G.L.Cavalleri, H.Hakonarson, E.L.Heinzen, M.R.Johnson, R.Kalviainen, B.P.C.Koeleman, R.Krause, P.Kwan, D.Lal, H.Lerche, Q.S.Li, I.Lopes-Cendes, D.H.Lowenstein, T.J.O'Brien, S.M.Sisodiya.

**Phenotyping committee:** C.Depondt, D.J.Dlugos, W.S.Kunz, P.Kwan, D.H.Lowenstein (Chair), A.G.Marson, P.Striano.

**Governance committee:** S.F.Berkovic, A.Compston, A-E.Lehesjoki, D.H.Lowenstein.

**Patient recruitment and phenotyping:** Z.Afawi, E.Amadori, A.Anderson, J.Anderson, D.M.Andrade, G.Annesi, A.Avbersek, M.D.Baker, G.Balagura, S.Balestrini, C.Barba, K.Barboza, F.Bartolomei, T.Bast, T.Baumgartner, B.Baykan, N.Bebek, A.J.Becker, F.Becker, C.A.Bennett, B.Berghuis, S.F.Berkovic, A.Beydoun, C.Bianchini, F.Bisulli, I.Blatt, I.Borggraefe, C.Bosselmann, V.Braatz, K.Brockmann, R.J.Buono, R.M.Busch, H.Caglayan, E.Campbell, L.Canafoglia, C.Canavati, G.D.Cascino, B.Castellotti, C.B.Catarino, F.Chassoux, K.Chinthapalli, I-J.Chou, S-K.Chung, P.O.Clark, A.J.Cole, A.Coppola, M.Cosico, P.Cossette, J.J.Craig, L.K.Davis, G-J.deHaan, N.Delanty, C.Depondt, P.Derambure, O.Devinsky, L.Di Vito, D.J.Dlugos, V.Doccini, C.P.Doherty, H.El-Naggar, C.E.Elger, C.A.Ellis, A.Faucon, L.Ferguson, T.N.Ferraro, L.Ferri, M.Feucht, M.Fitzgerald, B.Fonferko-Shadrach, F.Fortunato, S.Franceschetti, J.A.French, E.Freri, M.Gagliardi, A.Gambardella, E.B.Geller, T.Giangregorio, L.Gjerstad, T.Glauser, E.Goldberg, A.Goldman, T.Granata, D.A.Greenberg, R.Guerrini, K.Hallmann, M.Hegde, I.Helbig, C.Hengsbach, S.Hirose, E.Hirsh, H.Hjalgrim, P-C.Hung, M.Iacomino, L.L.Imbach, Y.Inoue, A.Ishii, J.Jamnadas-Khoda, L.Jehi, M.R.Johnson, R.Kälviainen, M.Kanaan, A.-M.Kantanen, B.Kara, S.M.Kariuki, D. Kasteleijn-Nolst Trenite, J.Kegele, Y.Kesim, N.Khoueiry-Zgheib, C.King, H.E.Kirsch, K.M.Klein, G. Kluger, S.Knake, R.C.Knowlton, A.D.Korczyn, A.Koupparis, I.Kousiappa, M.Krenn, H.Krestel, I.Krey, W.S.Kunz, G.Kurlemann, Ru.Kuzniecky, P.Kwan, A.Labate, A.Lacey, S.Lauxmann, S.L.Leech, A-E.Lehesjoki, J.R.Lemke, H.Lerche, G.Lesca, B.Neubauer, N.Lewin, Q.S.Li, L.Licchetta, K-L.Lin, D.Lindhout, T.Linnankivi, I.Lopes-Cendes, D.H.Lowenstein, C.H.T.Lui, F.Madia, A.G.Marson, C.M.McGraw, D.Mei, R.Minardi, R.S.Moller, M.Montomoli, B.Mostacci, L.Muccioli, H.Muhle, K.Müller-Schlüter, I.M.Najm, W.Nasreddine, C.R.J.C.Newton, T.J.O'Brien, Ç.Özkara, S.S.Papacostas, E.Parrini, M.Pendziwiat, W.O.Pickrell, R.Pinsky, T.Pippucci, An.Poduri, F.Pondrelli, R.H.W.Powell, M.Privitera, A.Rademacher, R.Radtke, F.Ragona, S.Rau, M.I.Rees, B.M.Regan, P.S.Reif, S.Rhelms, A.Riva, F.Rosenow, P.Ryvlin, A.Saarela, L.G.Sadleir, J.W.Sander, Th.Sander, M.Scala, Th.Scattergood, S.C.Schachter, C.J.Schankin, I.E.Scheffer, B.Schmitz, S.Schoch, S.Schubert-Bast, A.Schulze-Bonhage, P.Scudieri, B.R.Sheidley, J.J.Shih, G.J.Sills, S.M.Sisodiya, M.C.Smith, P.E.Smith, A.C.M.Sonsma, M.R.Sperling, B.J.Steinhoff, U.Stephani, W.C.Stewart, C.Stipa, P.Striano, H.Stroink, A.Strzelczyk, R.Surges, T.Suzuki,

K.M.Tan, G.A.Tanteles, E.Tauboll, L.L.Thio, O.Timonen, P.Tinuper, M.Todaro, P.Topaloglu, R.Tozzi, M-H.Tsai, B.Tumiene, D.Turkdogan, A.Utkus, P.Vaidiswaran, L.Valton, A.van Baalen, A.Vetro, E.P.G.Vining, F.Visscher, S.von Brauchitsch, R.von Wrede, R.G.Wagner, Y.G.Weber, S.Weckhuysen, J.Weisenberg, M.Weller, C.D.Whelan, P.Widdess-Walsh, M.Wolff, S.Wolking, D.Wu, K.Yamakawa, Z.Yapici, E.Yücesan, S.Zagaglia, F.Zahnert, F.Zimprich, G.Zsurka, Q.Zulfiqar Ali.

**Control cohorts:** L.C.Brody, J.G.Eriksson, A.Franke, H.Hakonarson, Y.-L. Lau, J.L.Mills, A.M.Molloy, M.M.Nöthen, A.Palotie, F.Pangilinan, H.Stroink, W.Yang.

**Consortium coordination:** K.L.Oliver.

# References


1. Fisher, R. S. *et al.* ILAE official report: a practical clinical definition of epilepsy. *Epilepsia* **55**, 475–482 (2014).
2. Fiest, K. M. *et al.* Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. *Neurology* **88**, 296–303 (2017).
3. 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.* **9**, 5269 (2018).
4. Epi4K consortium & Epilepsy Phenome/Genome Project. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol.* **16**, 135–143 (2017).
5. Leu, C. *et al.* Polygenic burden in focal and generalized epilepsies. *Brain* **142**, 3473–3481 (2019).
6. Vadlamudi, L. *et al.* Timing of de novo mutagenesis--a twin study of sodium-channel mutations. *N. Engl. J. Med.* **363**, 1335–1340 (2010).
7. Speed, D. *et al.* Describing the genetic architecture of epilepsy through heritability analysis. *Brain* **137**, 2680–2689 (2014).
8. Motelow, J. E. *et al.* Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. *Am. J. Hum. Genet.* **108**, 965–982 (2021).
9. Chen, Z., Brodie, M. J., Liew, D. & Kwan, P. Treatment Outcomes in Patients With Newly Diagnosed Epilepsy Treated With Established and New Antiepileptic Drugs: A 30-Year Longitudinal Cohort Study. *JAMA Neurol.* **75**, 279–286 (2018).
10. Devinsky, O. *et al.* Epilepsy. *Nat Rev Dis Primers* **4**, 18024 (2018).
11. Chan, S. S. L. & Copeland, W. C. DNA polymerase gamma and mitochondrial disease: understanding the consequence of POLG mutations. *Biochim. Biophys. Acta* **1787**, 312–319 (2009).
12. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012).
13. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
14. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
15. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
16. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
17. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
18. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).

19. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, (2018).
20. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
21. Xu, C. *et al.* Knockdown of RMI1 impairs DNA repair under DNA replication stress. *Biochem. Biophys. Res. Commun.* **494**, 158–164 (2017).
22. International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address: epilepsy-austin@unimelb.edu.au. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **13**, 893–903 (2014).
23. Yoshida, M. *et al.* Identification of novel BCL11A variants in patients with epileptic encephalopathy: Expanding the phenotypic spectrum. *Clin. Genet.* **93**, 368–373 (2018).
24. Holland, D. *et al.* Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet.* **16**, e1008612 (2020).
25. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
26. Ruano, D. *et al.* Functional gene group analysis reveals a role of synaptic heterotrimeric G proteins in cognitive ability. *Am. J. Hum. Genet.* **86**, 113–125 (2010).
27. Lukyanetz, E. A., Shkryl, V. M. & Kostyuk, P. G. Selective blockade of N-type calcium channels by levetiracetam. *Epilepsia* **43**, 9–18 (2002).
28. Wang, S. J., Huang, C. C., Hsu, K. S., Tsai, J. J. & Gean, P. W. Inhibition of N-type calcium currents by lamotrigine in rat amygdalar neurones. *Neuroreport* **7**, 3037–3040 (1996).
29. Marson, A. *et al.* The SANAD II study of the effectiveness and cost-effectiveness of levetiracetam, zonisamide, or lamotrigine for newly diagnosed focal epilepsy: an open-label, non-inferiority, multicentre, phase 4, randomised controlled trial. *Lancet* **397**, 1363–1374 (2021).
30. Christensen, J., Kjeldsen, M. J., Andersen, H., Friis, M. L. & Sidenius, P. Gender differences in epilepsy. *Epilepsia* **46**, 956–960 (2005).
31. Magi, R., Lindgren, C. M. & Morris, A. P. Meta-analysis of sex-specific genome-wide association studies. *Genet. Epidemiol.* **34**, 846–853 (2010).
32. Gaborit, N. *et al.* Gender-related differences in ion-channel and transporter subunit expression in non-diseased human hearts. *J. Mol. Cell. Cardiol.* **49**, 639–646 (2010).
33. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
34. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417 (2019).

35. Jeste, S. S. & Tuchman, R. Autism Spectrum Disorder and Epilepsy: Two Sides of the Same Coin? *J. Child Neurol.* **30**, 1963–1971 (2015).

36. Long, S. *et al.* The Clinical and Genetic Features of Co-occurring Epilepsy and Autism Spectrum Disorder in Chinese Children. *Front. Neurol.* **10**, 505 (2019).

37. Sanders, S. J. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).

38. Lauxmann, S. *et al.* An SCN2A mutation in a family with infantile seizures from Madagascar reveals an increased subthreshold Na(+) current. *Epilepsia* **54**, e117–21 (2013).

39. Ben-Shalom, R. *et al.* Opposing Effects on NaV1.2 Function Underlie Differences Between SCN2A Variants Observed in Individuals With Autism Spectrum Disorder or Infantile Seizures. *Biol. Psychiatry* **82**, 224–232 (2017).

40. Mirza, N. *et al.* Using common genetic variants to find drugs for common epilepsies. *Brain Commun* **3**, fcab287 (2021).

41. Bourgeois, B. F. D. Chronic management of seizures in the syndromes of idiopathic generalized epilepsy. *Epilepsia* **44 Suppl 2**, 27–32 (2003).

42. Marson, A. G. *et al.* The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *Lancet* **369**, 1016–1026 (2007).

43. Punetha, J. *et al.* Biallelic CACNA2D2 variants in epileptic encephalopathy and cerebellar atrophy. *Ann Clin Transl Neurol* **6**, 1395–1406 (2019).

44. Fariello, R. G. Safinamide. *Neurotherapeutics* **4**, 110–116 (2007).

45. Alsaegh, H., Eweis, H., Kamal, F. & Alrafiah, A. Celecoxib Decrease Seizures Susceptibility in a Rat Model of Inflammation by Inhibiting HMGB1 Translocation. *Pharmaceuticals* **14**, (2021).

46. Ma, M.-G. *et al.* RYR2 Mutations Are Associated With Benign Epilepsy of Childhood With Centrotemporal Spikes With or Without Arrhythmia. *Front. Neurosci.* **15**, 629610 (2021).

47. Yap, S. M. & Smyth, S. Ryanodine receptor 2 (RYR2) mutation: A potentially novel neurocardiac calcium channelopathy manifesting as primary generalised epilepsy. *Seizure* **67**, 11–14 (2019).

48. Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat. Genet.* **50**, 1593–1599 (2018).

49. Beesley, L. J. *et al.* The emerging landscape of health research based on biobanks linked to electronic health records: Existing resources, statistical challenges, and potential opportunities. *Stat. Med.* **39**, 773–800 (2020).

50. Wightman, D. P. *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, 1276–1282 (2021).

51. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, (2019).

52. Hautakangas, H. *et al.* Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat. Genet.* **54**, 152–160 (2022).

53. Reay, W. R. & Cairns, M. J. Advancing the use of genome-wide association studies for drug repurposing. *Nat. Rev. Genet.* **22**, 658–671 (2021).

54. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).

55. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).

56. de Bakker, P. I. W. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122–8 (2008).

57. Bhattacharjee, S. *et al.* A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am. J. Hum. Genet.* **90**, 821–835 (2012).

58. Cook, S. *et al.* Accurate imputation of human leukocyte antigens with CookHLA. *Nat. Commun.* **12**, 1264 (2021).

59. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

60. Choi, W., Luo, Y., Raychaudhuri, S. & Han, B. HATK: HLA analysis toolkit. *Bioinformatics* **37**, 416–418 (2021).

61. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

62. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

63. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).

64. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

65. de Klein, N. *et al.* Brain expression quantitative trait locus and network analysis reveals downstream effects and putative drivers for brain-related diseases. *bioRxiv* 2021.03.01.433439 (2021) doi:10.1101/2021.03.01.433439.

66. Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).

67. Weeks, E. M. *et al.* Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *bioRxiv* (2020) doi:10.1101/2020.09.08.20190561.

68. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).

69. Freytag, S., Burgess, R., Oliver, K. L. & Bahlo, M. brain-coX: investigating and visualising gene co-expression in seven human brain transcriptomic datasets. *Genome Med.* **9**, 55 (2017).

70. Rodriguez-Acevedo, A. J., Gordon, L. G., Waddell, N., Hollway, G. & Vadlamudi, L. Developing a gene panel for pharmacoresistant epilepsy: a review of epilepsy pharmacogenetics. *Pharmacogenomics* **22**, 225–234 (2021).

71. Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).

72. Speed, D., Holmes, J. & Balding, D. J. Evaluating and improving heritability models using summary statistics. *Nat. Genet.* **52**, 458–462 (2020).

73. Grotzinger, A. D., de la Fuente, J., Nivard, M. G. & Tucker-Drob, E. M. Pervasive downward bias in estimates of liability scale heritability in GWAS meta-analysis: A simple solution. *bioRxiv* (2021) doi:10.1101/2021.09.22.21263909.

74. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, (2018).

75. Zhong, S. *et al.* A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018).

76. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

77. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).

78. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).

79. Sudlow, C. *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.* **12**, e1001779 (2015).

80. Nagai, A. *et al.* Overview of the BioBank Japan Project: Study design and profile. *J. Epidemiol.* **27**, S2–S8 (2017).

81. Locke, A. E. *et al.* Exome sequencing of Finnish isolates enhances rare-variant association power. *Nature* **572**, 323–328 (2019).

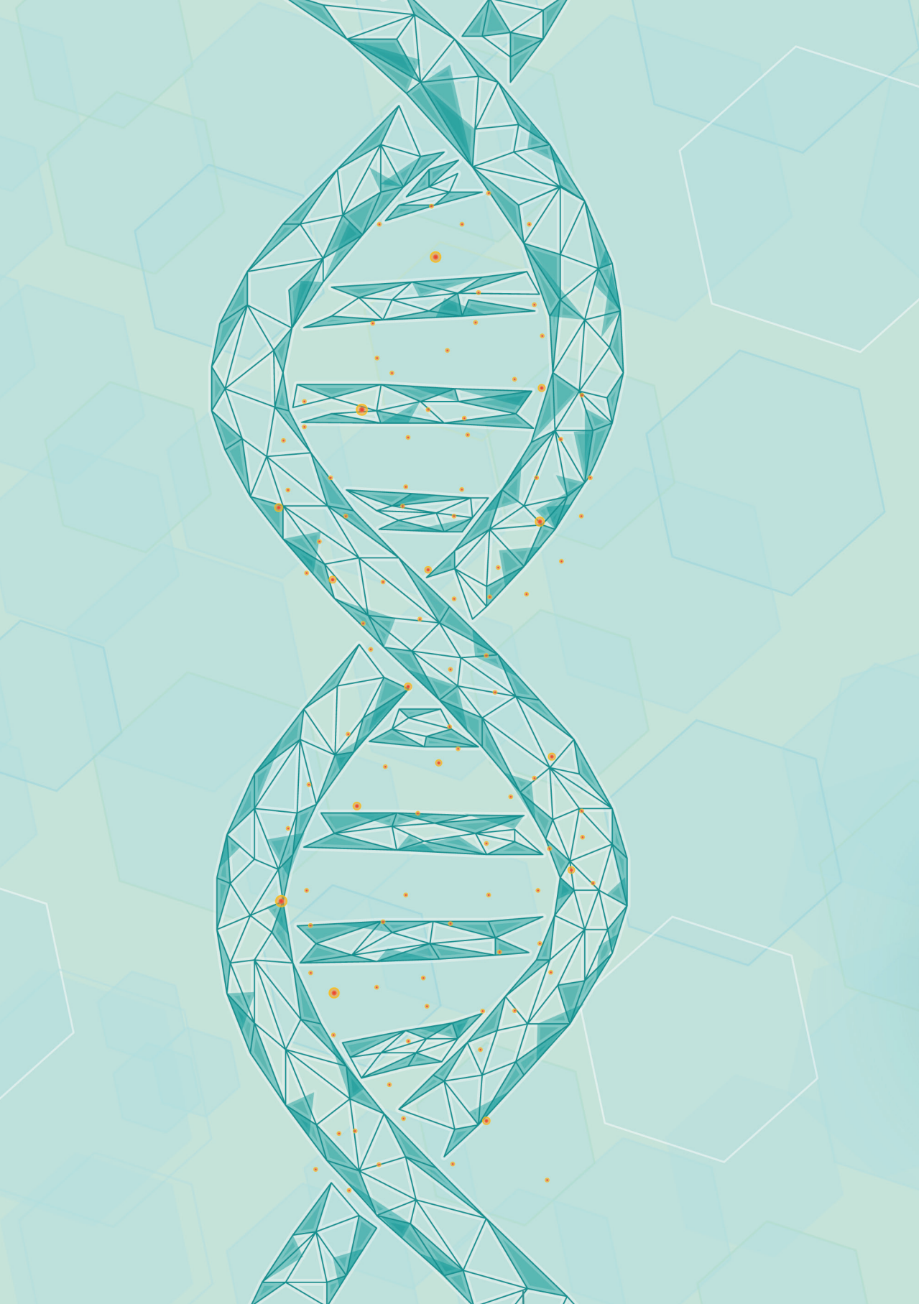
82. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).

83. Xue, A. et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat. Commun*. **9**, 2941 (2018).

84. Hautakangas, H. et al. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat. Genet*. **54**, 152–160 (2022).

85. Wood, M. D. & Gillard, M. Evidence for a differential interaction of brivaracetam and levetiracetam with the synaptic vesicle 2A protein. *Epilepsia* **58**, 255–262 (2017).

86. Bhattacharjee, S. et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am. J. Hum. Genet.* **90**, 821–835 (2012).



# CHAPTER 7

# DISTINCT GENETIC BASIS OF COMMON EPILEPSIES AND STRUCTURAL MRI MEASURES

Remi Stevelink, Bobby P.C. Koeleman, Sanjay M. Sisodiya, International League against Epilepsy Consortium on Complex epilepsies

*Manuscript under review*

## Summary

Focal and generalized epilepsies are associated with robust differences in MRI measures of subcortical structures, grey matter and white matter. However, it is unknown whether such structural brain differences reflect the cause or consequence of epilepsy or its treatment. Analyses of common genetic variants underlying both epilepsy and variability in structural brain measures can give further insights, since such inherited variants are not influenced by disease or treatment. Here, we performed genetic correlation analyses using data from the largest genome-wide association study (GWAS) on epilepsy (n=27,559 cases and 42,436 controls) and GWAS on MRI measures of white (n=33,292) or grey matter (n=51,665). We did not detect any significant genetic correlation between any type of epilepsy and any of 280 measures of grey matter, white matter or subcortical structures. These results suggest that there is a distinct genetic basis underlying risk of epilepsy and structural brain measures. This would imply that the genetic basis of normal structural brain variation is unrelated to that of common epilepsy. Structural changes in epilepsy could rather be the consequence of epilepsy, its comorbidities or its treatment, offering a cumulative record of disease.

# Main text

## Introduction

Large-scale collaborative efforts by the ENIGMA-Epilepsy working group have found widespread structural brain differences in people with generalized as well as focal epilepsy, when compared to healthy controls. These differences often extend considerably beyond any localised epileptogenic focus in the brain.[1,2] Such structural brain differences are thought to underlie various traits like cognitive decline and vulnerability to psychiatric diseases,[3] that are frequently comorbid in people with epilepsy. The ENIGMA-Epilepsy studies are based on cross-sectional comparisons of MRI scans between people with epilepsy and healthy controls, which do not allow for inference of causation. Therefore, it is unknown whether such structural brain differences constitute the cause of epilepsy, the result of epileptic seizures or epiphenomena such as effects of anti-epileptic drugs or environmental factors, or some combination of such factors.



Some of these limitations can be overcome by assessing common genetic factors associated with structural brain measurements and genetic factors associated with epilepsy in independent cohorts. Common inherited genetic variants are not determined by disease, treatment or environmental factors. Susceptibility to epilepsy and variation in structural brain measures are both strongly heritable and largely explained by common genetic variation.[3–5] Recent large-scale efforts have combined genetic and MRI data from tens of thousands of people, to map which genetic variants are associated with structural measures of the brain.[3,5] Combining this data with a large-scale GWAS from the epilepsies represents a unique opportunity to disentangle whether the genetic basis of epilepsy and structural brain variation are shared or distinct.

To do so, we performed genetic correlation analyses to assess whether genetic determinants of structural brain measures are associated with epilepsy and its main subtypes, focal and generalized epilepsy.

# Methods

## Study population

The current study is based on summary statistics from the International League Against Epilepsy Consortium on Complex Epilepsies epilepsy GWAS and structural MRI GWAS from the ENIGMA consortium and the UK Biobank. The epilepsy GWAS constitutes an unpublished meta-analysis combining previously published data[4] (n=14,534 people with epilepsy and 24,218 subjects without epilepsy) with unpublished GWAS data from the Epi25 collaborative[6] (n=13,025 people with epilepsy and 18,218 controls). In total, 27,559 people with epilepsy and 42,436 controls were included in the epilepsy meta-analysis. Furthermore, sub-analyses were conducted on focal (n=14,939 cases) and generalized epilepsy (n=6,952 cases). All included subjects were of European ancestry.

We used publicly available summary statistics from the ENIGMA consortium GWAS of 70 measures of grey matter, calculated from genetic data and brain MRI scans in 51,665 individuals of primarily (94%) European ancestry.[3] These 70 measures consisted of cortical thickness and surface area of 34 brain regions as well as the total surface area and average thickness of the whole cortex. The ENIGMA GWAS constitutes a meta-analysis involving 60 different cohorts, including various population-based cohorts like the UK Biobank as well as case-control cohorts, including a cohort of 178 subjects with epilepsy. Analyses were corrected for disease status for case-control cohorts.

White matter microstructure GWAS data was obtained from a study which combined genetic data with 110 measures of diffusion-weighted brain MRI scans from the UK Biobank in 34,024 subjects of European (British) ancestry.[5] These 110 measures consist of five diffusion tensor imaging (DTI) parameters (fractional anisotropy, mean diffusivity, axial diffusivity, radial diffusivity and mode of anisotropy) computed for 21 white matter tracts plus a whole brain average. Total brain volume data was obtained from a GWAS which included 47,316 subjects of European ancestry.[7] Furthermore, GWAS data using MRI scans of 19,629 subjects of European ancestry was used to assess genetic contribution to 100 brain volumetric phenotypes,

including various subcortical, cortical and white matter volumes. These subjects were primarily derived from the UK Biobank, plus a couple of smaller population-based cohorts.

### Genetic correlation analyses

Genetic correlation analyses between epilepsy and MRI measures of grey and white matter were computed using bivariate linkage disequilibrium score regression (LDSC).[8] We were not able to exclude any potential sample overlap. However, genetic correlations computed by LDSC are not biased by sample overlap.[8] We used default settings of LDSC, with precomputed linkage disequilibrium regression weights from European subjects of the 1000 Genomes project. We computed all genetic correlations analyses separately for all epilepsy combined and its main subtypes focal and generalized epilepsy.



### Power calculations

We used the GCTA-GREML power calculator[9] to estimate the power to detect significant genetic correlations of 0.05 and 0.10 or higher (at a type 1 error rate ($\alpha$) of 0.05) between all epilepsy, focal epilepsy, generalized epilepsy and each main MRI phenotype of the whole brain. SNP based heritability for each phenotype was calculated using LDSC[8] and converted to liability scale.[4] For these calculations, we assumed a population prevalence of 0.005 for all epilepsy, 0.003 for focal epilepsy and 0.002 for generalized epilepsy.[4]

## Results

We did not find any even nominally significant genetic correlation (all $p>0.05$) between all epilepsy, focal or generalized epilepsy with average surface area, cortical thickness, brain volume, fractional anisotropy, mean diffusivity, axial diffusivity, radial diffusivity or mode of anisotropy of the whole brain (Figure 1). Power calculations showed that we had statistical power ranging between 25% and 100% to detect genetic correlations higher than 0.05 between any of the epilepsy subtypes and any of the MRI phenotypes (Supplementary table 1). The power to detect genetic correlations

higher than 0.10 ranged between 73% and 100%. For generalized epilepsy (the epilepsy subtype with the highest SNP-based heritability), we had >95% power to detect any genetic correlation higher than 0.05 and 100% power to detect any genetic correlation higher than 0.10.

Next, we assessed the correlation between genetics of epilepsy subtypes with cortical thickness and surface area in 34 cortical brain regions, region brain volumes in 100 brain areas as well as five DTI measures in 21 white matter tracts (Supplementary table 2). These analyses yielded only 21 nominally significant genetic correlations amongst 819 tests (lowest p=0.002); none of these were significant when correcting for multiple testing (all p>0.05/819).

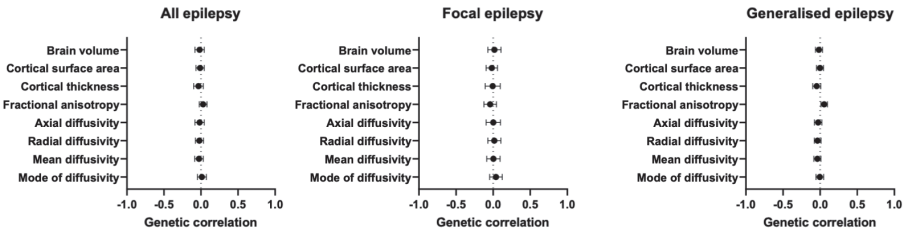



**Figure 1: No genetic correlation between epilepsy and structural measures of brain volume, grey matter or white matter of the whole brain.** Genetic correlations were computed using LDSC and the genetic correlation coefficients (±standard error) are plotted on the X-axis .

# Discussion

Here, we utilized the largest available GWASs to assess the genetic correlation between epilepsy and structural brain measures. We did not find any genetic overlap between epilepsy and any measure of grey or white matter of the brain. These results suggest that the genetic basis of epilepsy is distinct from the genetic basis of normal structural brain variation.

Previous studies that compared MRI scans between controls and people with focal and generalized epilepsy showed widespread, as well as regional and syndrome-specific, structural brain differences in white and grey matter of the brain.[1,2] It is known that the vulnerability for epilepsy, in particular generalized epilepsy, as well as the variance in grey and white matter MRI

measures, is substantially explained by common genetic variants.[3–5] The absence of genetic overlap between epilepsy and MRI measures suggests that the genetic variation underlying structural brain differences between people does not meaningfully influence risk of epilepsy. Conversely, the findings also suggest that common genetic variants underlying susceptibility to epilepsy do not affect grey or white matter structural variation. Importantly, the lack of formal correlation may suggest that measured structural brain differences found in people with epilepsy are unrelated to the underlying genetic cause of the disease, and represent a separate source of information about epilepsy at group and individual levels. If so, structural brain differences are more likely a consequence of epilepsy or its treatment than being its cause. For example, seizure activity could cause progressive brain atrophy.[10] Although frequent seizures are associated with more pronounced atrophy,[11] such atrophy is also found in patients who have become seizure-free.[10] A recent study found that cortical thinning in epilepsy is mediated by microglial activation.[12] Furthermore, transient depletion of activated microglia did not affect seizures, but did prevent cortical thinning, suggesting that these processes are distinct and potentially modifiable. Alternatively, treatment of epilepsy by anti-epileptic drugs could also affect grey and white matter volume. For example, valproic acid, but not other anti-epileptic drugs, has been associated with smaller grey and white matter volumes.[13]



Our study should be considered in light of some limitations. In this study, we only assessed common genetic variants (defined as a minor allele frequency >1%) in common types of epilepsy. We cannot rule out the possibility that rare genetic variants or copy number variants contribute to both epilepsy risk and variation in brain measures, including during development. Indeed, it is well known that some rare variants causing developmental epileptic encephalopathies (DEE), and other epilepsies, are associated with gross structural brain abnormalities, whilst some genes implicated in DEE have roles during brain development.[14] Similarly, focal epileptogenic lesions can be caused by rare genetic variants.[15] Since we only assessed genetics, we cannot rule out the influence of environmental or epigenetic factors that influence both epilepsy and brain structure. Although we used the currently largest available GWAS of epilepsy and MRI measures, their sample sizes are still

relatively modest compared to GWAS of more readily available phenotypes such BMI or height. Our study is large enough to exclude a large genetic correlation. However, we cannot exclude the possibility of a small genetic overlap between epilepsy and structural brain measures were a larger GWAS to be tested. Our analyses are based on epilepsy GWAS split into three broad categories. We did not have access to sufficiently powered GWAS of more resolved epilepsy subtypes; therefore, we are unable to rule out whether there are genetic correlations between specific epilepsy subtypes (like mesial temporal lobe epilepsy) and structural brain measures. The brain MRI GWAS that we used for our analyses were primarily based on population-based cohorts including <1% people with epilepsy. Therefore, we cannot exclude the possibility that there are epilepsy-specific genetic variants that influence both structural MRI measures as well as epilepsy risk.

Altogether, our results suggest that common epilepsies and structural brain variation have a distinct genetic basis. These results could aid in understanding the pathophysiology of epilepsy and associated structural brain changes. If structural brain changes in common epilepsy are indeed the consequence of epilepsy rather than the cause, it would suggest that it is modifiable or even preventable. Potentially, preventing structural brain changes in epilepsy could reduce risk of comorbid psychiatric disorders or cognitive decline.

## Acknowledgements

We would like to thank the Ming Fund for providing funding for R.S. We are grateful to the people with epilepsy and volunteers who participated in this research. We thank the following clinicians and research scientists for their contribution through sample collection (cases and controls), data analysis, and project support: Geka Ackerhans, Muna Alwaidh, R. E. Appleton, Willem Frans Arts, Guiliano Avanzini, Paul Boon, Sarah Borror, Kees Braun, Oebele Brouwer, Hans Carpay, Karen Carter, Peter Cleland, Oliver C. Cockerell, Paul Cooper, Celia Cramp, Emily de los Reyes, Chris French, Catharine Freyer, William Gallentine, Michel Georges, Peter Goulding, Micheline Gravel, Rhian Gwilliam, Lori Hamiwka, Steven J. Howell, Adrian Hughes, Aatif Husain, Monica Islam, Floor Jansen, Mary Karn, Mark Kellett, Ditte B. Kjelgaard, Karl Martin Klein, Donna Kring, Annie W. C. Kung, Mark Lawden, Jo Ellen Lee,

Benjamin Legros, Leanne Lehwald, Edouard Louis, Colin H. T. Lui, Zelko Matkovic, Jennifer McKinney, Brendan McLean, Mohamad Mikati, Bethanie Morgan-Followell, Wim Van Paesschen, Anup Patel, Manuela Pendziwiat, Marcus Reuber, Richard Roberts, Guy Rouleau, Cathy Schumer, B. Sharack, Kevin Shianna, N. C. Sin, Saurabh Sinha, Laurel Slaughter, Sally Steward, Deborah Terry, Chang-Yong Tsao, T. H. Tsoi, Patrick Tugendhaft, Jaime-Dawn Twanow, Jorge Vidaurre, Sarah Weckhuysen, Pedro Weisleder, Kathleen White, Virginia Wong, Raju Yerra, Jacqueline Yinger, and all contributing clinicians from the Department of Clinical and Experimental Epilepsy at the National Hospital for Neurology and Neurosurgery and University College London Institute of Neurology. This work was in part supported by a Translational Research Scholars award from the Health Research Board of Ireland (Christopher D. Whelan) and by research grants from Science Foundation Ireland (16/RC/3948 and 13/CDA/2223), and cofunded under the European Regional Development Fund and by FutureNeuro industry partners. Further funding sources include Wellcome Trust (grant 084730); Epilepsy Society, UK, National Institute for Health Research (NIHR; 08-08-SCC); GIHE, National Institutes of Health (NIH) R01-NS-49306-01 (Russell J. Buono); NIH R01-NS-053998 (Daniel H. Lowenstein); GSCFE, NIH R01-NS-064154-01 (Russell J. Buono, Hakon Hakonarson); NIH UL1TR001070, Development Fund from the Children's Hospital of Philadelphia (Hakon Hakonarson); National Health and Medical Research Council program grant 1091593 (Samuel F. Berkovic, Ingrid E. Scheffer, Karen L. Oliver, Katja E. Boysen); Royal Melbourne Hospital Foundation Lottery Grant (Slavé Petrovski); Royal Melbourne Hospital Neuroscience Foundation (Terence J. O'Brien); European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreements 279062 (EpiPGX) and 602102, Department of Health NIHR Biomedical Research Centres funding scheme, European Community (EC; FP6 project EPICURE: LSHM-CT2006-037315); German Research Foundation (DFG; SA434/4-1/4-26-1 (Thomas Sander), WE4896/3-1); EuroEPINOMICS Consortium (European Science Foundation/DFG: SA434/5-1, NU50/8-1, LE1030/11-1, HE5415/3-1 [Thomas Sander, Peter Nürnberg, Holger Lerche, Ingo Helbig], RO 3396/2- 1); German Federal Ministry of Education and Research, National Genome Research Network (NGFNplus/EMINet: 01GS08120, and 01GS08123 [Thomas Sander, Holger Lerche]; IntenC, TUR 09/

I10 [Thomas Sander]); Netherlands National Epilepsy Fund (grant 04-08); EC (FP7 project EpiPGX 279062); and Research Grants Council of the Hong Kong Special Administrative Region, China project numbers HKU7623/08 M (Stacey S. Cherny, Patrick Kwan, Larry Baum, Pak C. Sham), HKU7747/ 07 M (Stacey S. Cherny., Pak C. Sham), and CUHK4466/06 M (Patrick Kwan, Larry Baum). Collection of Belgian cases was supported by the Fonds National de la Recherche Scientifique, Fondation Erasme, Université Libre de Bruxelles. GlaxoSmithKline funded the recruitment and data collection for the GenEpA Consortium samples. We acknowledge the support of Nationwide Children's Hospital in Columbus, Ohio, USA. The Wellcome Trust (WT066056) and the NIHR Biomedical Research Centres Scheme (P31753) supported UK contributions. Further support was received through the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (contract N01HD33348). The project was also supported by the popgen 2.0 network through a grant from the German Ministry for Education and Research (01EY1103). Parts of the analysis of this work were performed on resources of the High Performance Center of the University of Luxembourg and Elixir-Luxembourg. The KORA study was initiated and financed by the Helmholtz Zentrum München–German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences, Ludwig Maximilian University, as part of LMUinnovativ. The ILAE facilitated the Consortium on Complex Epilepsies through the Commission on Genetics and by financial support; however, the opinions expressed in the article do not necessarily represent the policy or position of the ILAE.

## Disclosure of Conflicts of Interest

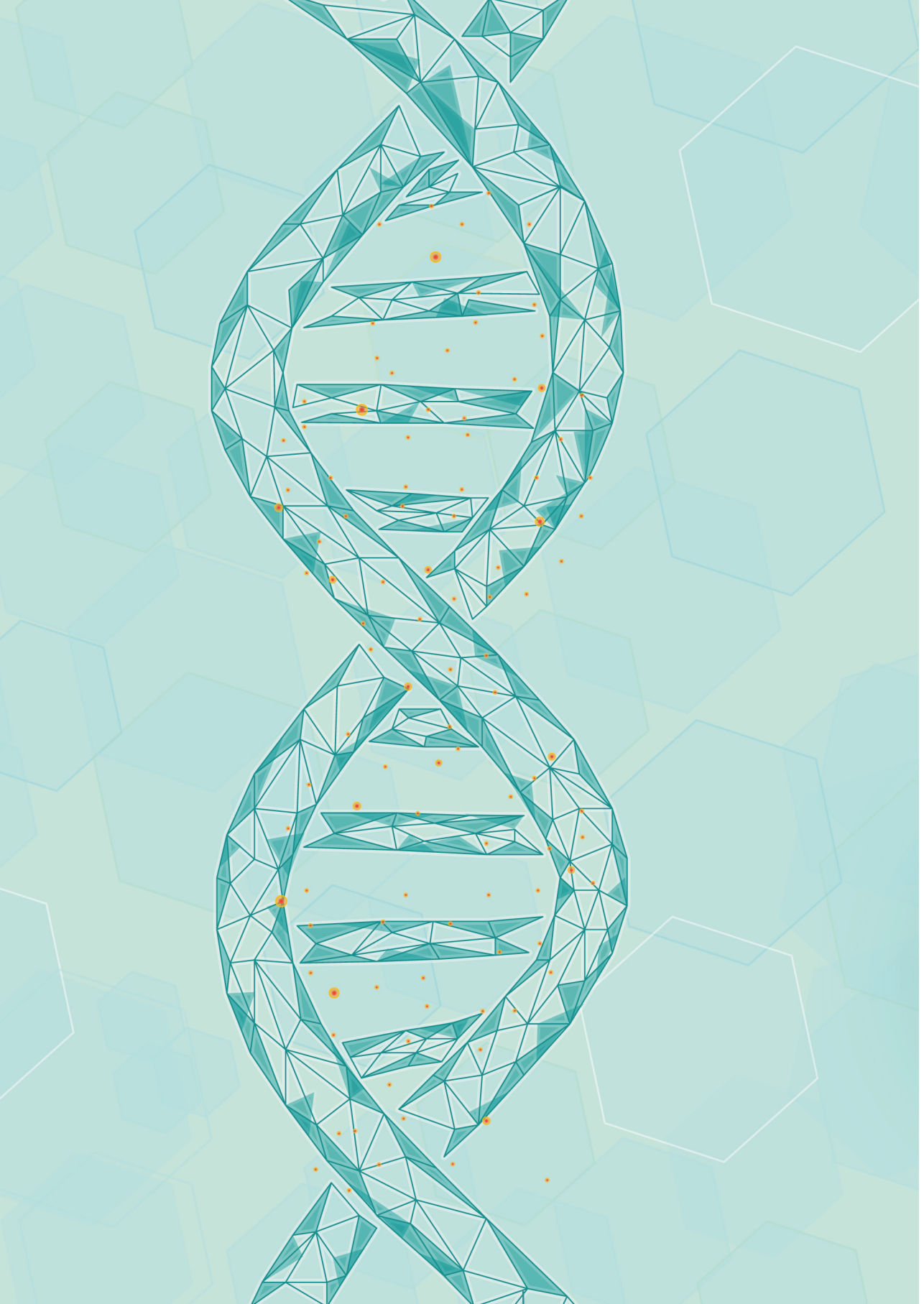
None of the authors has any conflict of interest to disclose.

## Supporting information

Supporting information can be found at: https://tinyurl.com/4j3srm5j

# References


1. Hatton SN, Huynh KH, Bonilha L, Abela E, Alhusaini S, Altmann A, et al. White matter abnormalities across different epilepsy syndromes in adults: an ENIGMA-Epilepsy study. Brain. 2020; 143(8):2454–73.
2. Whelan CD, Altmann A, Botía JA, Jahanshad N, Hibar DP, Absil J, et al. Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. Brain. 2018; 141(2):391–408.
3. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, et al. The genetic architecture of the human cerebral cortex. Science [Internet]. 2020; 367(6484). Available from: http://dx.doi.org/10.1126/science.aay6690
4. 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(1):5269.
5. Zhao B, Li T, Yang Y, Wang X, Luo T, Shan Y, et al. Common genetic variation influencing human white matter microstructure. Science [Internet]. 2021; 372(6548). Available from: http://dx.doi.org/10.1126/science.abf3736
6. Motelow JE, Povysil G, Dhindsa RS, Stanley KE, Allen AS, Feng Y-CA, et al. Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. Am J Hum Genet. 2021; 108(6):965–82.
7. Jansen PR, Nagel M, Watanabe K, Wei Y, Savage JE, de Leeuw CA, et al. Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. Nat Commun. 2020; 11(1):5606.
8. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015; 47(11):1236–41.
9. Visscher PM, Hemani G, Vinkhuyzen AAE, Chen G-B, Lee SH, Wray NR, et al. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. PLoS Genet. 2014; 10(4):e1004269.
10. Galovic M, van Dooren VQH, Postma TS, Vos SB, Caciagli L, Borzì G, et al. Progressive Cortical Thinning in Patients With Focal Epilepsy. JAMA Neurol. 2019; 76(10):1230–9.
11. Coan AC, Campos BM, Yasuda CL, Kubota BY, Bergo FP, Guerreiro CA, et al. Frequent seizures are associated with a network of gray matter atrophy in temporal lobe epilepsy with or without hippocampal sclerosis. PLoS One. 2014; 9(1):e85843.
12. Altmann A, Ryten M, Di Nunzio M, Ravizza T, Tolomeo D, Reynolds RH, et al. A systems-level analysis highlights microglial activation as a modifying factor in common epilepsies. Neuropathol Appl Neurobiol [Internet]. 2021; . Available from: http://dx.doi.org/10.1111/nan.12758
13. Tondelli M, Vaudano AE, Sisodiya SM, Meletti S. Valproate Use Is Associated With Posterior Cortical Thinning and Ventricular Enlargement in Epilepsy Patients. Front Neurol. 2020; 11:622.
14. Smith RS, Walsh CA. Ion Channel Functions in Early Brain Development. Trends Neurosci. 2020; 43(2):103–14.
15. Weckhuysen S, Marsan E, Lambrecq V, Marchal C, Morin-Brureau M, An-Gourfinkel I, et al. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. Epilepsia. 2016; 57(6):994–1003.



## CHAPTER 8

# POLYGENIC BURDEN IN FOCAL AND GENERALIZED EPILEPSIES

Costin Leu, Remi Stevelink, Alexander W 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

*Brain*. 2019; Nov 1;142(11):3473–3481.

## Abstract

Rare genetic variants can cause epilepsy, and genetic testing has been widely adopted for severe, paediatric-onset epilepsies. The phenotypic consequences of common genetic risk burden for epilepsies and their potential future clinical applications have not yet been determined. Using polygenic risk scores (PRS) from a European-ancestry genome-wide association study in generalized and focal epilepsy, we quantified common genetic burden in patients with generalized epilepsy (GE-PRS) or focal epilepsy (FE-PRS) from two independent non-Finnish European cohorts (Epi25 Consortium, $n = 5705$; Cleveland Clinic Epilepsy Center, $n = 620$; both compared to 20 435 controls). One Finnish-ancestry population isolate (Finnish-ancestry Epi25, $n = 449$; compared to 1559 controls), two European-ancestry biobanks (UK Biobank, $n = 383\ 656$; Vanderbilt biorepository, $n = 49\ 494$), and one Japanese-ancestry biobank (BioBank Japan, $n = 168\ 680$) were used for additional replications. Across 8386 patients with epilepsy and 622 212 population controls, we found and replicated significantly higher GE-PRS in patients with generalized epilepsy of European-ancestry compared to patients with focal epilepsy (Epi25: $P = 1.64\times10^{-15}$; Cleveland: $P = 2.85\times10^{-4}$; Finnish-ancestry Epi25: $P = 1.80\times10^{-4}$) or population controls (Epi25: $P = 2.35\times10^{-70}$; Cleveland: $P = 1.43\times10^{-7}$; Finnish-ancestry Epi25: $P = 3.11\times10^{-4}$; UK Biobank and Vanderbilt biorepository meta-analysis: $P = 7.99\times10^{-4}$). FE-PRS were significantly higher in patients with focal epilepsy compared to controls in the non-Finnish, non-biobank cohorts (Epi25: $P = 5.74\times10^{-19}$; Cleveland: $P = 1.69\times10^{-6}$). European ancestry-derived PRS did not predict generalized epilepsy or focal epilepsy in Japanese-ancestry individuals. Finally, we observed a significant 4.6-fold and a 4.5-fold enrichment of patients with generalized epilepsy compared to controls in the top 0.5% highest GE-PRS of the two non-Finnish European cohorts (Epi25: $P = 2.60\times10^{-15}$; Cleveland: $P = 1.39\times10^{-2}$). We conclude that common variant risk associated with epilepsy is significantly enriched in multiple cohorts of patients with epilepsy compared to controls—in particular for generalized epilepsy. As sample sizes and PRS accuracy continue to increase with further common variant discovery, PRS could complement established clinical biomarkers and augment genetic testing for patient classification, comorbidity research, and potentially targeted treatment.

# Introduction

Epilepsy is a common chronic neurological disorder, affecting approximately 1% of individuals (Ngugi *et al.,* 2010). Lifetime prevalence is 8–10% for a seizure and 3–4% for epilepsy (Hesdorffer *et al.,* 2011). The median incidence of epilepsy is 50 per 100 000 person-years (Ngugi *et al.,* 2011). Individuals at high risk for recurrent seizures (epilepsy) benefit from early antiseizure drug treatment, compared to no treatment or delayed treatment (Kim *et al.,* 2006). Predicting whether an individual will develop epilepsy after the first epileptic seizure is difficult (MacDonald *et al.,* 2000; Bell *et al.,* 2016), with recurrence risk varying from 27% to 71% (Hopkins *et al.,* 1988; Berg and Shinnar, 1991; Kwan and Sander, 2004).

Epileptic seizures either have a generalized (involving both cerebral hemispheres) or a focal (originating from one cerebral hemisphere) onset (Scheffer *et al.,* 2017). Generalized epilepsies account on average for 54%, focal epilepsies for 40%, and unclassifiable epilepsies for 7% of incident epilepsies in population-based studies of all ages (Banerjee *et al.,* 2009). Distinguishing between the two types of epilepsy can be difficult: focal epilepsy can present with bilateral tonic-clonic seizures (secondary-generalization), patients with generalized epilepsy can have focal features on EEG (Japaridze *et al.,* 2016), and some individuals have a mix of focal and generalized epilepsy (Scheffer *et al.,* 2017). Since commonly used antiseizure drugs for focal epilepsy can be ineffective or exacerbate generalized epilepsies, differentiating between focal and generalized epilepsy is important (Japaridze *et al.,* 2016). Hence, there is a clinical need for biomarkers that can help to distinguish individuals at high versus low risk to develop either focal or generalized epilepsy.

Genetic factors can explain a substantial portion of cases of epilepsy, particularly severe epilepsy (EpiPM Consortium *et al.,* 2015). For rare and early onset childhood epilepsies, >100 epilepsy-related genes have been discovered in recent years (Heyne *et al.,* 2019). The identified genetic variants are rare and of large effect, ranging from large deletions that confer on average ~7-fold risk for epilepsy (Pérez-Palma *et al.,* 2017) to single, causative *de novo* variants in >33 genes (Heyne *et al.,* 2018). These

variants are diagnostically relevant and can influence patient management. For example, treatment with sodium channel blockers can exacerbate seizures in patients with Dravet syndrome or other early-onset epileptic syndromes caused by *SCN1A* mutations, whereas these drugs are beneficial in patients with gain-of-function variants in *SCN2A* (Guerrini *et al.,* 1998; Löscher, 2009; Wolff *et al.,* 2017). While rare variation of large effect has a clear impact in clinical practice for rare epilepsy syndromes (McTague *et al.,* 2016), patients affected by common types of epilepsy rarely carry such variants and routine genetic testing is therefore not established for the common epilepsies.

Genome-wide association studies (GWAS) for common forms of epilepsy have identified common genetic risk variants for generalized epilepsy, focal epilepsy, and febrile seizures (Kasperavičiūtė *et al.,* 2013; Feenstra *et al.,* 2014; International League Against Epilepsy Consortium on Complex Epilepsies, 2014, 2018). Common genetic risk variants associated with a disease are usually of small effect size (1.33 median odds ratio) (Hindorff *et al.,* 2009) and cannot individually quantify risk or to inform prognosis and treatment. However, polygenic risk scores (PRSs) that combine the effect sizes of thousands of variants into a single score can stratify affected and healthy individuals. For five common disorders, a recent study showed a 3- to 5-fold increased risk for patients with a high disease-specific PRS, similar to the range of risk conferred by rare monogenic variants, such as *LDLR* missense variants for coronary artery disease or rare *BRCA* variants in breast cancer (Khera *et al.,* 2018). Based on these results, the authors proposed that PRS-based prediction may be reliable enough to consider their utility in clinical practice.

Genome-wide PRS based on thousands of common variants associated with epilepsy may help distinguish healthy individuals from those who develop epilepsy (Speed *et al.,* 2014). However, no studies have directly investigated whether PRSs derived from well-phenotyped cohorts stratify patients in clinical practice or population-based cohort studies. Here, we calculate PRSs for the two main subtypes of epilepsy (generalized and focal) from the largest GWAS in epilepsy to date (International League Against Epilepsy Consortium on Complex Epilepsies, 2018) and (i) quantify the burden of

PRSs derived from GWAS studies of well-phenotyped cohorts in patients with generalized or focal epilepsy; (ii) explore if PRS can differentiate patients with generalized from those with focal epilepsy; and (iii) explore if patients with generalized or focal epilepsy are enriched particularly in the upper extreme of the PRS burden distribution compared to controls. Our overall study design is presented in Fig. 1. Across two independent research cohorts, one clinically ascertained cohort, and three biobanks (repositories with clinical data and DNA samples available for research); data from 630 598 individuals were available for the PRS analyses.

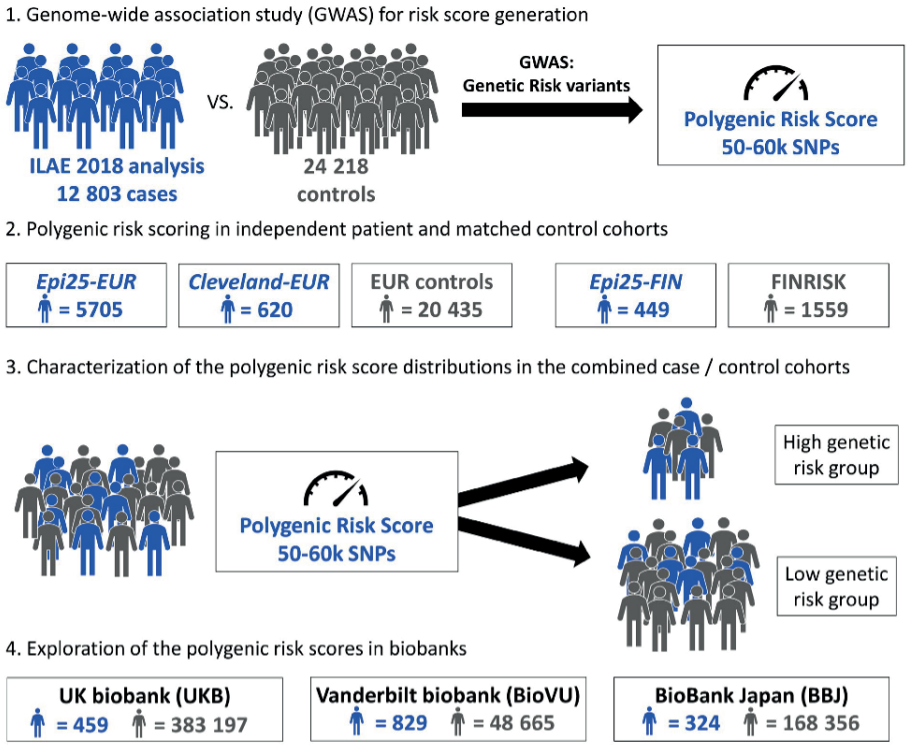



**Figure 1: Study design.** (1) PRSs for the two main classes of epilepsy (generalized and focal) were derived from the largest GWAS in epilepsy to date (International League Against Epilepsy Consortium on Complex Epilepsies, 2018). (2) PRS were calculated in patients with generalized or focal epilepsy and in population controls and (3) tested in their ability to identify significant differences of common variant burden among groups. (4) The UK and Vanderbilt biobanks were available to test the behaviour of the PRSs in individuals ascertained by ICD-10 codes for epilepsy, while the Biobank Japan was available to test the performance in a non-European population. SNP = single nucleotide polymorphism.

# Patients and methods

## Study cohorts

Patients of European ancestry with generalized epilepsy or focal epilepsy were recruited through the Epi25 project (http://epi-25.org/), an international multicentre epilepsy genetics research consortium [exploration cohorts for generalized (GE) and focal (FE) epilepsy: GE-Epi25-EUR and FE-Epi25-EUR, respectively] and from a single clinical centre, the Cleveland Clinic Epilepsy Center (replication cohorts for generalized and focal epilepsy: GE-Cleveland-EUR and FE-Cleveland-EUR, respectively). A Finnish-ancestry population isolate was recruited from the Epi25 project (GE-Epi25-FIN and FE-Epi25-FIN, respectively). Ancestry-matched population controls were recruited from several in-house projects, the Partners HealthCare Biobank (Karlson *et al.,* 2016), and the FINRISK study (Borodulin *et al.,* 2017). Three large-scale biobank repositories [UKB: UK Biobank (Sudlow *et al.,* 2015); BioVU: Vanderbilt University biorepository (Roden *et al.,* 2008); and BBJ: BioBank Japan (Nagai *et al.,* 2017)] were used for additional explorations. All cohorts, totalling 630 598 individuals, are detailed in Table 1 and the Supplementary material.

| Cohort name | Ascertainment type | Ethnicity | Generalized epilepsy (GE) | Focal epilepsy (FE) | Controls |
|---|---|---|---|---|---|
| Epi25-EUR | Research | EUR | 2256 | 3449 | 20 435 |
| Cleveland-EUR | Clinic | EUR | 85 | 535 | 20 435 |
| Epi25-FIN | Research | FIN | 112 | 337 | 1559 |
| UKB | Biobank | EUR | 246 | 213 | 383 197 |
| BioVU | Biobank | EUR | 293 | 536 | 48 665 |
| BBJ | Biobank | JPN | 219 | 105 | 168 356 |

**Table 1: Study cohorts after quality control.** Generalized and focal epilepsy were diagnosed in the Epi25-EUR, Cleveland-EUR, Epi25-FIN, and BBJ cohorts according to clinical criteria (clinical interview, neurological examination, EEG, imaging data). For the UK and BioVU biobanks, ICD-10 G40.3 codes were used to identify people with generalized epilepsy, and G40.0 to G40.2 codes to identify people with focal epilepsy. BBJ = BioBank Japan; BioVU = Vanderbilt University biorepository; Cleveland-EUR = European-ancestry Cleveland Clinic Epilepsy Center cohort; Epi25-EUR = European-ancestry Epi25 cohort; Epi25-FIN = Finnish-ancestry Epi25 cohort; UKB = UK Biobank.

## Polygenic risk scoring in the study cohorts

Single-nucleotide polymorphism (SNP) weights for PRS were derived from summary statistics of the ILAE Consortium on Complex Epilepsies GWAS for generalized and focal epilepsy (International League Against Epilepsy Consortium on Complex Epilepsies, 2018). SNP weights for negative control PRS were derived from the UKB GWAS for type 2 diabetes for all cohorts excluding the UKB, and from the DIAGRAM-type 2 diabetes GWAS (Scott *et al.*, 2017) for the UKB. PRS for each individual were generated using the allelic scoring function, as implemented in PLINK v1.9 (Chang *et al.*, 2015). Individual PRSs were calculated as the sum of weighted effect alleles divided by the number of SNPs in the analysis. We generated the PRSs at the *P*-value threshold 0.5, found to be the best predicting threshold in a random split (80% training, 20% validation) of our exploration cohort (Epi25-EUR, Supplementary material 4.8, Supplementary Tables 4 and 5). We excluded individuals if their data were included in the GWAS studies used for PRS development. Details of the method to detect overlapping individuals across cohorts and the SNP quality control applied are given in the Supplementary material.



## Statistical analysis

We used logistic regression adjusted for sex and for the first four principal components of ancestry to determine the ability of PRS to stratify cases from controls. The proportion of phenotypic variance explained by PRS was calculated using Nagelkerke's pseudo-$R^2$, by comparing the full model of the logistic regression (PRS plus all covariates: sex and the first four principal components of ancestry) to the null model (covariates only). Following the example of Khera *et al.* (2018) we assessed the enrichment of the two epilepsy phenotypes (generalized epilepsy or focal epilepsy) in progressively more extreme tails of the PRS distribution (top 20%, 5%, 0.5%) against the remainder of the distribution in a logistic regression model predicting disease status, 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).

Fixed-effect meta-analysis, with adjustment for the effective sample size, was performed using METAL (Willer *et al.*, 2010).

### Data availability

The data that support the findings of this study are available from the Epi25 Consortium, upon reasonable request. The biobank data are available from the UKB, BioVU, and BBJ upon successful project application.

## Results

### Higher PRS burden in patients with epilepsy compared to controls

To determine if common variants associated with epilepsy are enriched in independent cohorts of patients with generalized epilepsy or focal epilepsy compared to population controls, we conducted a PRS analysis in two independent epilepsy cohorts of European ancestry. We found that in the GE-Epi25-EUR cohort, genome-wide polygenic risk scores for generalized epilepsy (GE-PRS) were significantly higher in patients with generalized epilepsy ($n$ = 2256 cases) than in population controls ($n$ = 20 435 controls; $P = 2.35 \times 10^{-70}$; Fig. 2 and Supplementary Table 1). GE-PRS explained 2.8% of the total phenotypic variance (composed of genetic, environmental, and genetic-environmental interaction variances) among the case and control group of the GE-Epi25-EUR cohort (Supplementary Table 1). This observation was replicated in the clinical GE-Cleveland-EUR cohort ($P = 1.43\times10^{-7}$; $n$ = 85 cases; 2.6% phenotypic variance explained). In the FE-Epi25-EUR cohort, the genome-wide polygenic risk for focal epilepsy (FE-PRS) was significantly higher in patients with focal epilepsy ($n$ = 3449 cases) than in population controls ($P = 5.74\times10^{-19}$), with 0.6% of the phenotypic variance explained. This observation was replicated in the clinical FE-Cleveland-EUR cohort ($n$ = 535; $P = 1.69\times10^{-6}$; 0.5% phenotypic variance explained). As expected, PRSs for type 2 diabetes (negative control) were not significantly higher in patients with generalized epilepsy or focal epilepsy than in population controls (Fig. 2 and Supplementary Table 1).

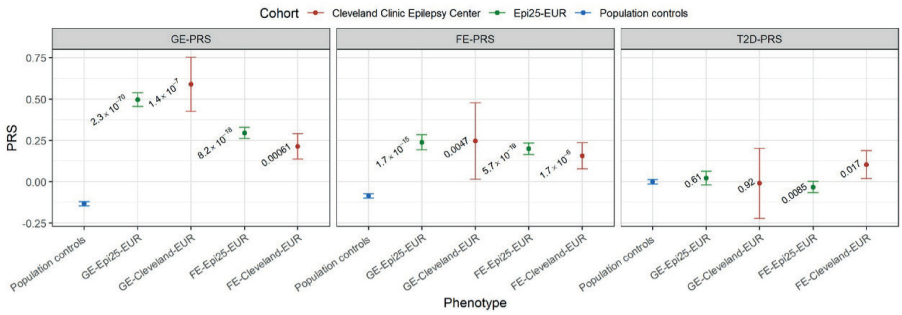



**Figure 2: Genome-wide polygenic risk for generalized epilepsy or focal epilepsy in the exploration and replication cohorts.** Shown are the means of the standardized GE-, FE-, and type 2 diabetes-PRS with 95% confidence intervals for the European-ancestry population controls (highlighted in blue; n = 20 435), the European-ancestry generalized epilepsy and focal epilepsy Epi25 exploration cohorts (highlighted in green; GE-Epi25-EUR, n = 2256; FE-Epi25-EUR, n = 3449), and the European-ancestry generalized epilepsy and focal epilepsy Cleveland Clinic replication cohorts (highlighted in red; GE-Cleveland-EUR, n = 85; FE-Cleveland-EUR, n = 535). 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 $\alpha = 1.67 \times 10\text{-}2$ (three tests per cohort).

To test the utility of PRSs across different populations, we investigated the power of the PRS derived from the European population in the isolated Finnish population. The GE-PRS was significantly higher in patients with generalized epilepsy ($n$ = 112 cases) than in the population controls ($n$ = 1559 controls; $P = 3.11 \times 10^{-4}$; Supplementary Fig. 2 and Supplementary Table 2). However, the PRSs explained less phenotypic variance than generalized epilepsy cohorts of European ancestry (2% phenotypic variance explained). The FE-PRS were not significantly different between Finnish patients with focal epilepsy ($n$ = 337 cases) and controls ($P$ = 0.55).

## Higher PRS burden in generalized compared to focal epilepsy

To determine if common variants associated with generalized epilepsy are enriched in patients with generalized epilepsy compared to patients with focal epilepsy, we regressed the GE-PRS against the diagnosis of generalized epilepsy or focal epilepsy. In the Epi25-EUR cohort, GE-PRSs were significantly higher in patients with generalized epilepsy than in those with focal epilepsy ($P = 1.64 \times 10^{-15}$), explaining 1.7% of the phenotypic variance (Supplementary Table 1). This observation was replicated in the Cleveland-

EUR cohort ($P$ = 2.85×10$^{-4}$, 3.9% phenotypic variance explained) and the Epi25-FIN cohort ($P$ = 1.80×10$^{-4}$, 4.6% phenotypic variance explained, Supplementary Table 2). Overall, the PRS had the most predictive power in the corresponding epilepsy phenotype group: GE-status was best predicted by GE-PRS ($P$ = 2.35×10$^{-70}$, 2.8% phenotypic variance explained in Epi25-EUR) over FE-PRS ($P$ = 1.71×10$^{-15}$, 0.6% phenotypic variance explained) and FE-status was best predicted by FE-PRS ($P$ = 5.74×10$^{-19}$, 0.6% phenotypic variance explained in Epi25-EUR) over GE-PRS ($P$ = 8.21×10$^{-18}$, 0.5% phenotypic variance explained, Supplementary Table 1).

## Enrichment of patients with epilepsy in the highest PRS burden percentile

To explore if the GE- and FE-PRS enrichment in patients with epilepsy is due to a few patients with a very high burden or due to many with a slightly elevated burden, we characterized the epilepsy PRS distribution in the European-ancestry cohorts. Patients with epilepsy and population control subjects were ranked according to their PRSs and tested for enrichment of patients with generalized epilepsy or focal epilepsy in the extreme tails of the PRS distribution. Strikingly, in the combined GE-Epi25-EUR and control cohorts, we observed a significant 4.63-fold enrichment of patients with generalized epilepsy in the group with the highest GE-PRS (top 0.5%, $P$ = 2.60×10$^{-15}$; involving 2.39% of the GE-Epi25-EUR cohort; Table 2). This observation was replicated in the clinical GE-Cleveland-EUR cohort (4.47-fold enrichment; $P$ = 1.39×10$^{-2}$; 3.53% of the GE-Cleveland-EUR cohort). In the FE-Epi25-EUR cohort, we observed a significant 2-fold enrichment of patients with focal epilepsy in the group with the highest FE-PRS (top 0.5%, $P$ = 5.57×10$^{-4}$; 1.16% of the FE-Epi25-EUR cohort; Table 2). This observation was not replicated in the smaller clinical FE-Cleveland-EUR cohort ($P$ = 0.22). All patients with top 0.5% highest PRS were found in the top decile of the GE- and FE-PRS distributions of the Epi25-EUR cohort (Supplementary Figs 3 and 4). Measures of the diagnostic accuracy of the PRS are given in Supplementary Table 8.

| | Reference group | OR | 95% CI | *P*-value | Cases/ controls upper PRS% | Cases/ controls lower PRS% | Sensitivity | Specificity |
|---|---|---|---|---|---|---|---|---|
| **GE–PRS / GE-Epi25** | | | | | | | | |
| Top 20% of distribution | Remaining 80% | 2.04 | 1.86–2.25 | $4.61 \times 10^{-48}$ | 887/3652 | 1369/16 783 | 0.393 | 0.821 |
| Top 5% of distribution | Remaining 95% | 2.39 | 2.06–2.76 | $4.39 \times 10^{-32}$ | 305/830 | 1951/19 605 | 0.135 | 0.959 |
| Top 0.5% of distribution | Remaining 99.5% | 4.63 | 3.16–6.76 | $2.60 \times 10^{-15}$ | 54/60 | 2202/20 375 | 0.024 | 0.997 |
| **GE–PRS / GE-Cleveland** | | | | | | | | |
| Top 20% of distribution | Remaining 80% | 2.09 | 1.33–3.28 | $1.31 \times 10^{-3}$ | 35/4070 | 50/16 365 | 0.412 | 0.801 |
| Top 5% of distribution | Remaining 95% | 2.02 | 1.00–3.72 | $3.44 \times 10^{-2}$ | 11/1016 | 74/19 419 | 0.129 | 0.950 |
| Top 0.5% of distribution | Remaining 99.5% | 4.47 | 1.07–12.6 | $1.39 \times 10^{-2}$ | 3/100 | 82/20 335 | 0.035 | 0.995 |
| **FE–PRS / FE-Epi25** | | | | | | | | |
| Top 20% of distribution | Remaining 80% | 1.35 | 1.24–1.48 | $2.32 \times 10^{-11}$ | 992/3785 | 2457/16 650 | 0.288 | 0.815 |
| Top 5% of distribution | Remaining 95% | 1.44 | 1.25–1.66 | $7.74 \times 10^{-7}$ | 292/903 | 3157/19 532 | 0.085 | 0.956 |
| Top 0.5% of distribution | Remaining 99.5% | 2.00 | 1.34–2.95 | $5.57 \times 10^{-4}$ | 40/80 | 3409/20 355 | 0.012 | 0.996 |
| **FE–PRS / FE-Cleveland** | | | | | | | | |
| Top 20% of distribution | Remaining 80% | 1.49 | 1.22–1.82 | $1.04 \times 10^{-4}$ | 148/4047 | 387/16 388 | 0.277 | 0.802 |
| Top 5% of distribution | Remaining 95% | 1.55 | 1.09–2.14 | $1.17 \times 10^{-2}$ | 40/1009 | 495/19 426 | 0.075 | 0.951 |
| Top 0.5% of distribution | Remaining 99.5% | 1.83 | 0.62–4.32 | 0.22 | 5/100 | 530/20 335 | 0.009 | 0.995 |

**Table 2: Odds ratios (OR) and P-values were calculated by comparing those within the top 0.5, 5, and 20% of the PRS distribution, to the remainder of the PRS distribution in 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). CI = confidence interval; EUR = European.

### Epilepsy PRSs have limited value in biobank-derived cohorts based on ICD-10 codes

To evaluate whether common epilepsy variants identified in well-phenotyped patients are enriched in patient cohorts ascertained from patient registries and biobanks, we extended our GE- and FE-PRS analyses to include three biobank-derived cohorts. In the UKB and BioVU biobanks, diagnosis was ascertained by ICD-10 codes for epilepsy and by a standardized questionnaire for the attending physicians in the BBJ biobank. Fixed-effect meta-analysis of the two European biobanks (UKB and BioVU), adjusted for the effective sample size (Willer *et al.,* 2010), revealed significantly higher GE-PRS in individuals coded as having generalized epilepsy than in population controls ($P = 7.99 \times 10^{-4}$, 539 patients with generalized epilepsy against 431 862 population controls). However, the PRS explained only very little of the phenotypic variance in each biobank (UKB: 0.12% variance explained; BioVU: 0.19%). In the BBJ, the GE-PRS were not significantly different between Japanese-ancestry patients with generalized epilepsy and controls ($P = 0.33$, 219 patients with generalized epilepsy against 168 356 controls). The FE-PRS were not significantly different between individuals coded as having focal epilepsy and controls in any of the three studied biobanks (UKB: $P = 0.44$; BioVU: $P = 0.23$; BBJ: $P = 0.29$). The PRSs for type 2 diabetes (negative control) were not significantly different in patients with generalized epilepsy or focal epilepsy than in population controls. The results in the three biobanks are detailed in Supplementary Table 3.

## Discussion

We identified a significantly higher genetic burden for epilepsy, as quantified by PRS, in independent cohorts of patients with epilepsy as compared to population controls. While modest effect sizes preclude risk prediction based on single common genetic variants, PRSs that combine thousands of variants show predictive ability across a range of complex traits and diseases, including neuropsychiatric disorders (Khera *et al.,* 2018). In the setting of a collaborative epilepsy genetics community, we demonstrate that available datasets have reached an adequate size to address the role of common genetic variants, each with small effect sizes, in large populations

of patients with epilepsy. In line with previous studies showing significant differences between the genetic architectures of generalized epilepsy and focal epilepsy (Speed *et al.,* 2014; International League Against Epilepsy Consortium on Complex Epilepsies, 2018), we also show that patients with generalized epilepsy have a significantly higher burden of common risk variants associated with generalized epilepsy than patients with focal epilepsy. The PRSs perform similarly in a multicentre research cohort and in an unselected—although much smaller—cohort ascertained through routine clinical practice in one single hospital. In contrast, significant, but small differences of PRS burden in large-scale biobanks provide evidence that ICD-10 epilepsy codes are not the best substitutes for precise clinical classifications by experts, despite our efforts to identify patients with generalized or focal epilepsy using stringent ICD code filtering. In line with recent evidence, we observe that PRSs derived from a European cohort have lower power when applied to populations of different genetic architecture, as observed in the cohorts of Finnish and Japanese ancestry (Martin *et al.,* 2017).



By evaluating the PRS distribution, we identify patients with an effect size of polygenic variants at group level that is comparable to those in other studies with rare variants of large effect. Among the group of patients with high GE-PRS (top 0.5% of cases and controls with the highest scores), we observe an enrichment of patients with generalized epilepsy similar to that seen among carriers of established genetic risk factors, such as copy number variations and *de novo* variants: the largest copy number variation burden study in epilepsy to date showed a 7.45-fold enrichment of patients with generalized epilepsy (2.78% of all patients with generalized epilepsy) among all hotspot copy number variations carriers (cases/controls) (Pérez-Palma *et al.,* 2017). The largest *de novo* variant study in neurodevelopmental disorders with epilepsy showed a 4.6-fold excess of *de novo* variants in known genes associated with developmental and epileptic encephalopathies in neurodevelopmental disorders with epilepsy when compared to those without epilepsy (Heyne *et al.,* 2018). In this study, we identify a 4.63-fold enrichment of patients with generalized epilepsy in the top 0.5% highest GE-PRS in the Epi25 exploration cohort (2.39% of patients with generalized

epilepsy) and a 4.47-fold enrichment of patients with generalized epilepsy in the top 0.5% highest GE-PRS in the clinical replication cohort (3.53% of patients with generalized epilepsy).

PRSs could have clinical implications for epilepsies because of their predictive power. Treatment with antiepileptic drugs after the first seizure has been debated, and the decision to start pharmacological treatment is usually based on relative risks, benefits, and lifestyle factors. After the first seizure, ~50% of individuals go on to have a second seizure within 3–5 years, with most recurrences occurring within the first year (Kho *et al.*, 2006; Wiebe *et al.*, 2008). Several factors can increase the risk of seizure recurrence, including abnormal results on neurological examination, brain imaging, or EEG, a family history of epilepsy, or a personal history of remote symptomatic seizures (Wiebe *et al.*, 2008). Patients at high risk for recurrence have been shown to benefit from immediate antiepileptic drug treatment after a first seizure compared to no treatment or delayed treatment (Kim *et al.*, 2006). For an individual with new-onset epilepsy, it is also critical to differentiate between a focal versus generalized epilepsy to inform the selection of the first-line antiseizure drug (Perucca *et al.*, 1998; Goldenberg, 2010). Differential diagnosis is especially challenging for focal epilepsy patients with secondary generalization or for those not found to have a relevant lesion on magnetic resonance imaging scans. A PRS indicating that a person is carrying an excessive amount of common risk variants for epilepsy or for generalized versus focal epilepsy could provide useful information for clinicians in deciding when to begin, and what type of treatment should be provided.

Future research should determine if and how well PRS can improve existing prediction models when combined with other factors, including established genetic risk factors of individually larger effect. Although our study represents the first of its kind in epilepsy, it needs to be replicated in a prospective setting. The predictive power of the PRS is determined by the genetic homogeneity of the GWAS from which the PRS is generated and that of the cohort to which it is applied. For epilepsy, a strong Eurocentric bias in the only available large scale GWAS is impeding PRS prediction in non-European individuals. Possible approaches to improve the predictive power

in the non-European population, such as including the target population in the training data (Márquez-Luna *et al.,* 2017), have been explored, but to realize the full potential of the PRS, greater population diversity must be prioritized in future GWAS studies (Martin *et al.,* 2019). It is possible that PRSs for focal epilepsies currently lack power because this group is genetically and phenotypically more heterogeneous than the group of generalized epilepsy (Speed *et al.,* 2014; International League Against Epilepsy Consortium on Complex Epilepsies, 2018). The clinical value of the PRS will also be limited by the low prevalence rates of epilepsy, leading to high negative predictive values, but low positive predictive values. To facilitate the implementation of PRS into clinical practice, additional research with larger, better-differentiated cohorts from different populations with well-characterized epilepsy phenotypes will be needed.

In summary, common polygenic variant burden for epilepsy can be measured and is differently distributed among patients with epilepsy and controls as well as between the two main epilepsy phenotypes (i.e. generalized and focal). PRS for epilepsies can provide physicians with an estimate of an individual's overall genetic risk for epilepsy that could aid in early diagnosis and targeted treatment in the future. In addition, a combination of rare and common variants that may predispose an individual to develop epilepsy provides a chance for more informative prediction tools that may lead to a paradigm shift from current practice in rare disorder genetics (presence or absence of a Mendelian, high-risk variant) to a liability threshold model that assumes for each individual a continuous liability composed of rare and common genetic risk variants.



## Funding

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.

## Competing interests

The authors report no competing interests.

## Supplementary material

Supplementary material is available at: https://tinyurl.com/4j3srm5j.

# References

Banerjee PN, Filippi D, Allen Hauser W. The descriptive epidemiology of epilepsy-a review. Epilepsy Res 2009; 85: 31–45.

Bell GS, Neligan A, Giavasi C, Keezer MR, Novy J, Peacock JL, et al. Outcome of seizures in the general population after 25 years: a prospective follow-up, observational cohort study. J Neurol Neurosurg Psychiatry 2016; 87: 843–50.

Berg AT, Shinnar S. The risk of seizure recurrence following a first unprovoked seizure: a quantitative review. Neurology 1991; 41: 965–72.

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.

EpiPM ConsortiumBerkovic SF, Scheffer IE, Petrou S, Delanty N, Dixon-Salazar TJ, et al. A roadmap for precision medicine in the epilepsies. Lancet Neurol 2015; 14: 1219–28.

Feenstra B, Pasternak B, Geller F, Carstensen L, Wang T, Huang F, et al. Common variants associated with general and MMR vaccine-related febrile seizures. Nat Genet 2014; 46: 1274–82.

Goldenberg MM. Overview of drugs used for epilepsy and seizures: etiology, diagnosis, and treatment. P T 2010; 35: 392–415.

Guerrini R, Dravet C, Genton P, Belmonte A, Kaminska A, Dulac O. Lamotrigine and seizure aggravation in severe myoclonic epilepsy. Epilepsia 1998; 39: 508–12.

Hesdorffer DC, Logroscino G, Benn EKT, Katri N, Cascino G, Hauser WA. Estimating risk for developing epilepsy: a population-based study in Rochester, Minnesota. Neurology 2011; 76: 23–7.

Heyne HO, Artomov M, Battke F, Bianchini C, Smith DR, Liebmann N, et al. Targeted gene sequencing in 6994 individuals with neurodevelopmental disorder with epilepsy. Genet Med 2019; doi: 10.1038/s41436-019-0531-0.

Heyne HO, Singh T, Stamberger H, Abou Jamra R, Caglayan H, Craiu D, et al. De novo variants in neurodevelopmental disorders with epilepsy. Nat Genet 2018; 50: 1048–53.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009; 106: 9362–67.

Hopkins A, Garman A, Clarke C. The first seizure in adult life. Value of clinical features, electroencephalography, and computerised tomographic scanning in prediction of seizure recurrence. Lancet 1988; 1: 721–6.

International League Against Epilepsy Consortium on Complex Epilepsies. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. Lancet Neurol 2014; 13: 893–903.

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.

Japaridze G, Kasradze S, Lomidze G, Zhizhiashvili L, Kvernadze D, Geladze K, et al. Focal EEG features and therapeutic response in patients with juvenile absence and myoclonic epilepsy. Clin Neurophysiol 2016; 127: 1182–7.

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: 2.

Kasperavičiūtė D, Catarino CB, Matarin M, Leu C, Novy J, Tostevin A, et al. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. Brain 2013; 136: 3140–50.

Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018; 50: 1219–24.

Kho LK, Lawn ND, Dunne JW, Linto J. First seizure presentation: do multiple seizures within 24 hours predict recurrence? Neurology 2006; 67: 1047–9.

Kim LG, Johnson TL, Marson AG, Chadwick DW; MRC MESS Study Group. Prediction of risk of seizure recurrence after a single seizure and early epilepsy: further results from the MESS trial. Lancet Neurol 2006; 5: 317–22.

Kwan P, Sander JW. The natural history of epilepsy: an epidemiological view. J Neurol Neurosurg Psychiatry 2004; 75: 1376–81.

Löscher W. Preclinical assessment of proconvulsant drug activity and its relevance for predicting adverse events in humans. Eur J Pharmacol 2009; 610: 1–11.

MacDonald BK, Johnson AL, Goodridge DM, Cockerell OC, Sander JW, Shorvon SD. Factors predicting prognosis of epilepsy after presentation with seizures. Ann Neurol 2000; 48: 833–41.

Márquez-Luna C, Loh P-R; South Asian Type 2 Diabetes (SAT2D) Consortium; SIGMA Type 2 Diabetes Consortium, Price AL. Multiethnic polygenic risk scores improve risk prediction in diverse populations. Genet Epidemiol 2017; 41: 811–23.

Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. Human Demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet 2017; 100: 635–49.

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 2019; 51: 584–91.

McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol 2016; 15: 304–16.

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: S2–8.

Ngugi AK, Bottomley C, Kleinschmidt I, Sander JW, Newton CR. Estimation of the burden of active and life-time epilepsy: a meta-analytic approach. Epilepsia 2010; 51: 883–90.

Ngugi AK, Kariuki SM, Bottomley C, Kleinschmidt I, Sander JW, Newton CR. Incidence of epilepsy: a systematic review and meta-analysis. Neurology 2011; 77: 1005–12.

Pérez-Palma E, Helbig I, Klein KM, Anttila V, Horn H, Reinthaler EM, et al. Heterogeneous contribution of microdeletions in the development of common generalised and focal epilepsies. J Med Genet 2017; 54: 598–606.

Perucca E, Gram L, Avanzini G, Dulac O. Antiepileptic drugs as a cause of worsening seizures. Epilepsia 1998; 39: 5–17.

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: 362–9.

Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. Epilepsia 2017; 58: 512–21.

Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An expanded genome-wide association study of type 2 diabetes in Europeans. Diabetes 2017; 66: 2888–902.

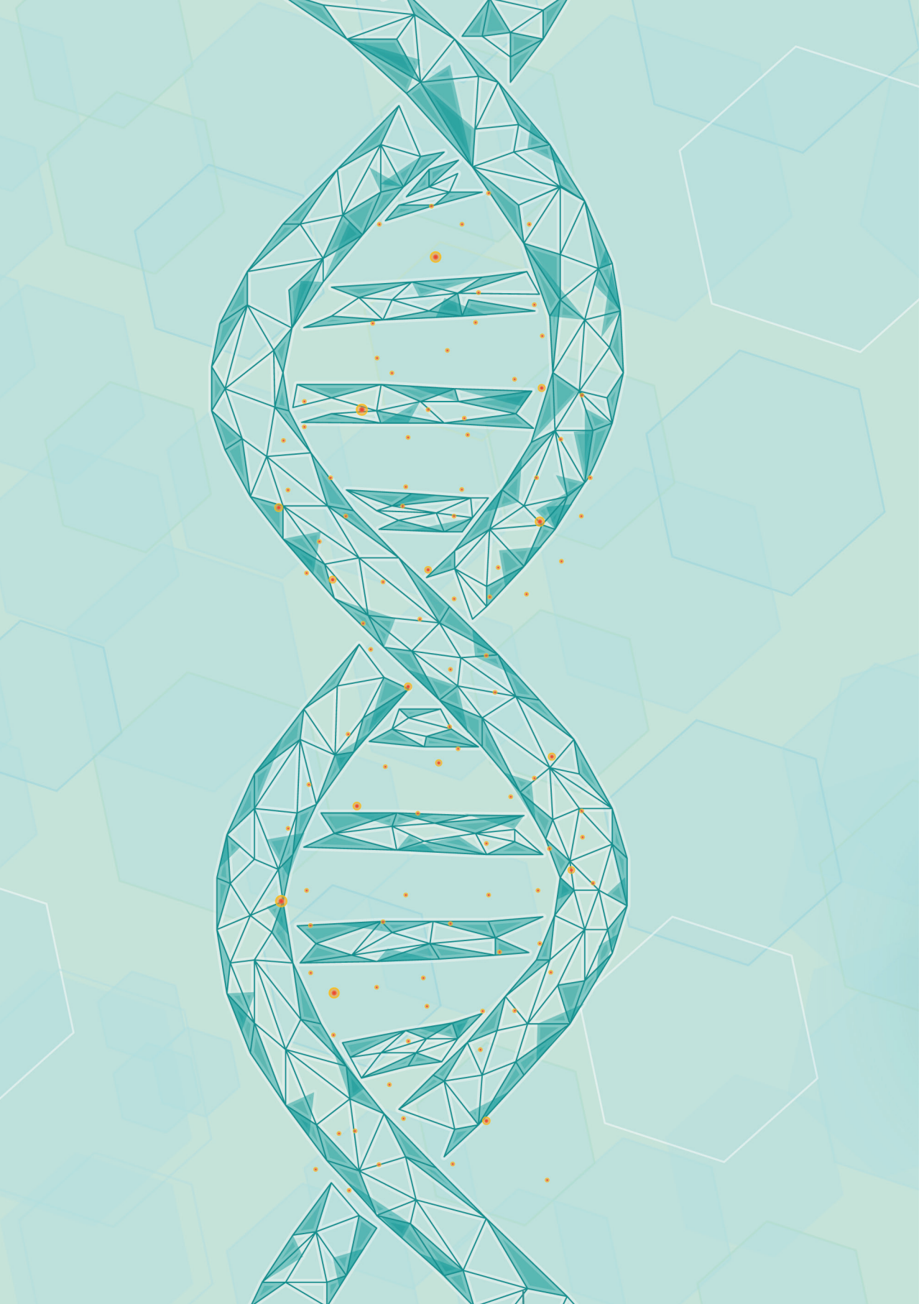
Speed D, O'Brien TJ, Palotie A, Shkura K, Marson AG, Balding DJ, et al. Describing the genetic architecture of epilepsy through heritability analysis. Brain 2014; 137: 2680–9.

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.

Wiebe S, Téllez-Zenteno JF, Shapiro M. An evidence-based approach to the first seizure. Epilepsia 2008; 49: 50–7.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26: 2190–1.

Wolff M, Johannesen KM, Hedrich UBS, Masnada S, Rubboli G, Gardella E, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. Brain 2017; 140: 1316–36.



## CHAPTER 9

# EPILEPSY SURGERY FOR PATIENTS WITH GENETIC REFRACTORY EPILEPSY: A SYSTEMATIC REVIEW

Remi Stevelink, Maurits W.C.B. Sanders, Maarten P. Tuinman, Eva H. Brilstra, Bobby P.C. Koeleman, Floor E. Jansen, Kees PJ. Braun

*Epileptic Disorders.* 2018; Apr 1;20(2):99-115.

# Abstract

**Aims**

In recent years, many different DNA mutations underlying the development of refractory epilepsy have been discovered. However, genetic diagnostics are still not routinely performed during presurgical evaluation and reports on epilepsy surgery outcome for patients with genetic refractory epilepsy are limited. We aimed to create an overview of the literature on seizure outcome following epilepsy surgery in patients with different genetic causes of refractory epilepsy.

**Methods**

We systematically searched PubMed and Embase prior to January 2017 and included studies describing treatment outcome following epilepsy surgery in patients with genetic causes of epilepsy. We excluded studies in which patients were described with epilepsy due to Tuberous Sclerosis Complex or Sturge-Weber syndrome (since this extensive body of research has recently been described elsewhere) and articles in which surgery was aimed to be palliative.

**Results**

We identified 24 eligible articles, comprising a total of 82 patients who had undergone surgery for (mainly childhood-onset) refractory epilepsy due to 15 different underlying genetic causes. The success rate of surgery varied widely across these different genetic causes. Surgery was almost never effective in patients with epilepsy due to mutations in genes involved in channel function and synaptic transmission, whereas surgery was significantly more successful regarding seizure control in patients with epilepsy due to mutations in the mTOR pathway. Patients with a lesion on MRI tended to have higher seizure freedom rates than those who were MRI-negative.

**Conclusion**

Although the evidence is still scarce, this systematic review suggests that studying genetic variations in patients with refractory epilepsy could help guide the selection of surgical candidates.

# Introduction

It is estimated that around 60% of epilepsy patients have focal epilepsy, of whom nearly half are medically refractory (West *et al.*, 2015). Epilepsy surgery is the only treatment that may be curative in patients with medically refractory epilepsy. However, epilepsy surgery is strongly under-utilised and currently less than half of refractory epilepsy patients are referred for evaluation of epilepsy surgery candidacy (de Flon *et al.*, 2010; Uijl *et al.*, 2012).

The relatively low proportion of potential surgical candidates who actually undergo surgery is largely due to a lack of factual information regarding epilepsy surgery and uncertainty around treatment outcome (Dewar and Pieters, 2015). Although there are several prognostic factors for surgical success (West *et al.*, 2015), it is often unclear for which patients surgery is indicated or contraindicated. Currently, on average, only 65% of patients achieve seizure freedom after surgery (West *et al.*, 2015).

Over recent years, it has increasingly been acknowledged that many patients with either generalized or focal types of epilepsy have an underlying genetic cause (Helbig *et al.*, 2008; Hildebrand *et al.*, 2013). These include single gene mutations that are related to channelopathies and disorders of synaptic transmission (Helbig *et al.*, 2008), or the mammalian target of rapamycin (mTOR) pathway, involved in various processes such as neuronal growth, migration, and proliferation (Baldassari *et al.*, 2016). In addition, there are several microdeletions and other chromosomal abnormalities that are known to be associated with epilepsy. This heterogeneity in molecular genetic aetiology points to differences in the underlying pathophysiology and is reflected by phenotypic differences between patients. It is possible that these different causes are also associated with differences in response to epilepsy surgery.

It is commonly accepted that epilepsy patients with a genetically determined focal structural lesion(s), such as tuberous sclerosis, may be candidates for surgery (Zhang *et al.*, 2013). However, many patients with genetic causes of epilepsy do not have detectable epileptogenic lesions on MRI, so called "MRI-negative" patients. In general, the absence of a visible brain lesion on MRI significantly decreases the chance of surgical success (Téllez-

Zenteno *et al.*, 2010; Bast, 2013). MRI-negative patients with focal epilepsy can still be considered surgical candidates (Bast *et al.*, 2016) as there may be an undetected underlying focal epileptogenic brain lesion, such as mild malformations of cortical development (mMCD) or focal cortical dysplasia (FCD) (So and Lee, 2014). Greater MR field strength, improved MRI sequences, and new post-processing techniques have increased the detection rate of such mMCDs and FCDs (So and Lee, 2014). Even in truly MRI-negative patients with refractory focal epilepsy and a consistent electrophysiological focus, epilepsy surgery is increasingly considered due to advances in multimodal functional neuroimaging and invasive monitoring techniques, such as stereo-electroencephalography (S-EEG). Pathology often subsequently reveals an underlying mMCD or FCD (So and Lee, 2014). However, surgery has been successful in some (18-47%) patients without demonstrated pathological abnormalities (Téllez-Zenteno *et al.*, 2010). A still unknown proportion of MRI-negative patients with focal refractory epilepsy who are evaluated for epilepsy surgery may have a genetic epilepsy syndrome. Identification of such genetic causes could have prognostic value for surgical outcome in these patients.

Genetic diagnostics are still not routinely performed in patients with refractory epilepsy, mostly due to the high costs and low throughput of traditional DNA sequencing techniques (Hildebrand *et al.*, 2013). The possibility to comprehensively test all epilepsy patients for genetic causes has been enhanced in recent years, with the advent of next-generation sequencing techniques (Hildebrand *et al.*, 2013).

To date, reports of epilepsy surgery for patients with genetic causes of epilepsy are sporadic. Some recent studies have shown that epilepsy surgery may be effective in patients with mutations in specific genes (Lee *et al.*, 2012; Jansen *et al.*, 2015), but this has never been shown in patients with other gene mutations (Barba *et al.*, 2014; Skjei *et al.*, 2015). Such findings suggest that routine genetic diagnostics for causative mutations of epilepsy prior to surgery could be of importance to determine surgical candidacy. This systematic review provides an overview of the reported outcomes of epilepsy surgery in patients with an established genetic cause of epilepsy. Future aims include the use of genetic diagnostics in the presurgical assessment

of patients with refractory epilepsy in order to assist the clinician in the often complex dilemma of whether to proceed to surgery or rather stop the time-consuming, costly, and often invasive, presurgical trajectory in patients with a proven genetic epilepsy syndrome.

# Methods

## Search strategy and study selection

Our search strategy and study selection are summarised in *figure 1*. A literature search in PubMed and Embase was performed by one author (RS) in order to identify articles in which epilepsy, genetics, and surgery were described together, using various synonyms (*supplementary tables S1 and S2*). The search was initially performed in June 2016 and updated in November 2016. The search yielded a total of 1,345 articles.

We included all studies reporting on epilepsy surgery and seizure outcome and collected details only on patients who either had a definite clinical diagnosis of a genetic syndrome with co-morbid epilepsy, or who had a mutation or other genetic abnormality detected that was highly likely to be the cause of their epilepsy. All patients with genetic causes of epilepsy who were described in the reports were included, regardless whether the causative mutation was somatic/mosaic or germline, although we describe the results for these subgroups separately. We excluded articles on patients with epilepsy due to tuberous sclerosis complex or Sturge Weber syndrome from this systematic review, since this extensive body of research has recently been described elsewhere (Bourgeois *et al.*, 2007; Zhang *et al.*, 2013). Furthermore, we excluded articles in which epilepsy surgery was described for patients with genetic mutations that were associated with, but not considered monogenic causes of, their epilepsy; for example, BRAF mutations in glioneuronal tumours, reported as potential prognostic factors for surgery outcome (Prabowo *et al.*, 2014). Moreover, we excluded surgical cases when the intention of surgery was stated to be palliative, rather than curative.

All search results were reviewed based on title and abstract. The full-text was reviewed in potentially eligible articles. Moreover, references of

the included articles were reviewed, as well as other articles in which the eligible articles are cited, using the "cited by" functions in PubMed and Embase. The article search and selection were checked by a second author (MS).

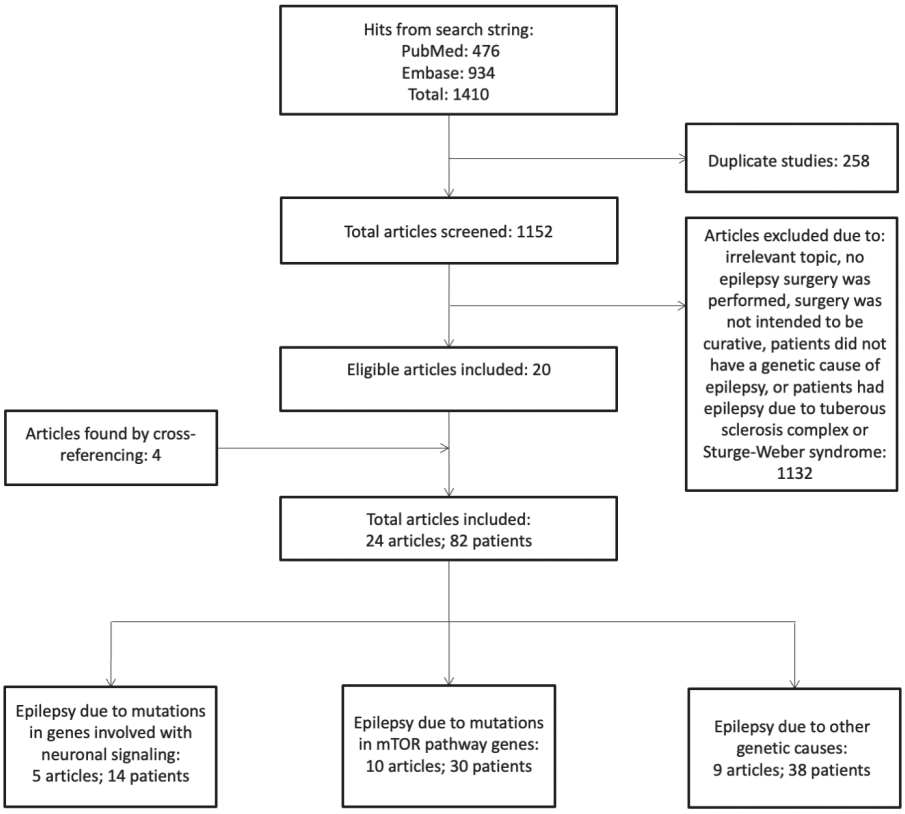



**Figure 1: Flowchart of search strategy and study selection.**

## Data processing

A standardised data extraction form was created, containing nine variables: affected gene, causative genetic variants, number of patients, histology of resected tissue, MRI findings, surgery type, mean follow-up time in years, post-surgical seizure outcome, and whether the surgery was successful. We divided the included articles into three main categories of genetic causes of epilepsy:

- Pathogenic variants of genes related to ion channel function and synaptic transmission.
- Pathogenic variants of mTOR pathway genes.
- Other genetic causes of epilepsy.

Extraction of raw data from the included articles was performed by RS and checked by MS.

Whenever possible, we classified histological descriptions of resected or isolated tissue according to the standardised classification system of focal cortical dysplasia (FCD) defined by the ILAE.

Where possible, we categorised descriptions of MRI findings as FCD, hippocampal sclerosis (HS) or hemimegalencephaly. Patients were defined as MRI-negative based on either no abnormalities or only non-specific abnormalities, not judged to be the cause of epilepsy, on MRI. All patients without detectable causative lesions on MRI were used for subgroup analysis.

Successful surgery was defined as Engel Class I ("free of disabling seizures"), the equivalent ILAE Class 1, or a description of seizure outcome equivalent to these classifications, based on the last reported follow-up visit.



# Results

## Search results

The literature search yielded a total of 20 eligible articles and a further four publications were identified through a cross-reference check of the citations of the included articles, as well as all publications in which the eligible articles are cited.

The 24 included studies described a total of 82 patients, with 15 different genetic causes of (mainly childhood-onset) epilepsy, who underwent surgery. The success rate of surgery varied widely amongst these different genetic causes (table 1).

## Genes related to channel function and synaptic transmission

The literature search yielded five articles that described a total of 14 surgery cases with epilepsy due to pathogenic variants in genes related to ion channel

function and synaptic transmission (table 2). These epileptogenic mutations were found in the voltage-gated sodium channels *SCN1A* and *SCN1B* (Helbig *et al.*, 2008), the gene *CNTNAP2* which is involved in AMPA-receptor trafficking and excitatory neuronal network activity (Anderson *et al.*, 2012; Varea *et al.*, 2015), and *STXBP1*, which is involved in the release of neurotransmitters (Weckhuysen *et al.*, 2013).

Epilepsy surgery did not lead to complete seizure freedom in any of the eight patients with *SCN1A* mutations who underwent epilepsy surgery, even though six of them had focal seizure semiology which co-localized with MRI-visible lesions (Barba *et al.*, 2014; Skjei *et al.*, 2015). Outcome data concerning specific seizure types that were primarily targeted by the surgical procedure (*e.g.* temporal lobe seizures in patients with HS) were not provided in the publications included. Seven of the patients with *SCN1A* mutation had a clinical phenotype consistent with Dravet syndrome and the other had a clinical phenotype most consistent with genetic epilepsy with febrile seizures plus (GEFS+). Two patients had no MRI-visible lesion.

Two patients underwent surgery for epilepsy due to mutations in *SCN1B* and both patients became seizure-free after temporal lobectomy (Scheffer *et al.*, 2007); one had underlying HS, whereas no brain abnormality, on MRI or histopathological examination of resected tissue, was detected in the other patient.

All three patients with epilepsy due to a homozygous mutation in *CNTNAP2* had a recurrence of seizures after surgery (Strauss *et al.*, 2006).

One patient with epilepsy due to a *STBXP1* mutation underwent surgery since she had prominent focal findings on EEG, despite having no abnormalities on MRI. Epilepsy surgery did not lead to cessation of seizures although her seizure frequency had decreased (Weckhuysen *et al.*, 2013). Pathology of the resected tissue revealed FCD.

Overall, surgery was successful regarding the control of seizures for only two of 14 patients (14%) with pathogenic variants in genes related to channelopathies and disorders of synaptic transmission.

| Genetic cause | | MRI-positiveseizure-free/total | MRI-negativeseizure-free/total | Total groupseizure-free/total |
|---|---|---|---|---|
| **Pathogenic variants of genes related to ion channel function and synaptic transmission** | *SCN1A* | FCD: 0/2HS: 0/2Encephalomalacia: 0/1Subcortical area of abnormal signal: 0/1 | 0/2 | 0/8 |
| | *SCN1B* | HS: 1/1 | 1/1 | 2/2 |
| | *CNTNAP2* | HS: 0/2 | 0/1 | 0/3 |
| | *STXBP1* | - | 0/1 | 0/1 |
| **Overall** | | 1/9 | 1/5 | 2/14 (14%) |
| **Pathogenic variants of mTOR pathway genes** | *DEPDC5* | FCD: 3/6 | 2/3 | 5/9 |
| | *PTEN* | HME: 1/1 | - | 1/1 |
| | *NPRL2* | - | 0/1 | 0/1 |
| | *NPRL3* | FCD: 1/1 | - | 1/1 |
| **Overall** | | 5/8 | 2/4 | 7/12 (58%) |
| **Other genetic causes of epilepsy** | Microdeletions | HS: 9/10 | 0/2 | 9/12 |
| | Neurofibromatosis type 1 | FCD: 2/2HS: 4/6Polymicrogyria: 0/1Tumour: 5/11 | 1/1 | 12/21 |
| | Fragile-X syndrome | HS: 2/2 | - | 2/2 |
| | Mitochondrial mutations | HS: 1/3 | - | 1/3 |
| **Overall** | | **23/35** | **1/3** | **24/38 (63%)** |
| **Total** | | **29/52 (56%)** | **4/12 (33%)** | **33/64(52%)** |

**Table 1A: Success rates of epilepsy surgery for patients with different genetic causes (germline mutations) of epilepsy.** FCD: focal cortical dysplasia; HS: hippocampal sclerosis. HME: hemimegalencephaly.

| Genetic cause | | MRI-positiveseizure-free/total | MRI-negativeseizure-free/total | Total groupseizure-free/total |
|---|---|---|---|---|
| Pathogenic variants of mTOR pathway genes | *PIK3CA* | HME: 5/5FCD: 1/1 | - | 6/6 |
| | *AKT3* | HME: 1/3FCD: 1/1 | - | 2/4 |
| | *mTOR* | HME: 1/1FCD: 6/7 | - | 7/8 |
| **Total** | | **15/18 (83%)** | **-** | **15/18 (83%)** |

**Table 1B: Success rates of epilepsy surgery for patients with different genetic causes (somatic mutations) of epilepsy.** FCD: focal cortical dysplasia; HME: hemimegalencephaly.

## mTOR pathway genes

The search yielded 10 articles that described a total of 30 patients who underwent surgery for epilepsy in relation to mutations in the following mTOR pathway genes: *DEPDC5*, *PTEN*, *PIK3CA*, *AKT3*, *NPRL2*, *NPRL3,* and mTOR itself (table 3). In 12 patients, germline mutations were found in *DEPDC5*, *PTEN*, *NPRL2* or *NPRL3* genes, whereas in 18 patients (somatic or mosaic) mutations were detected in resected tissue, involving the genes *PIK3CA*, *AKT3,* and mTOR.

Epilepsy surgery controlled seizures completely in seven of 12 patients with mutations in *DEPDC5*, *PTEN*, *NPRL2* or *NPRL3*, of whom eight had a lesion on MRI (Baulac *et al.,* 2015; Carvill *et al.,* 2015; Jansen *et al.,* 2015; Scerri *et al.,* 2015; Weckhuysen *et al.,* 2016). Three more patients had a significant improvement in seizure frequency, whereas two patients had no improvement.

Fifteen of 18 patients with somatic or mosaic mutations in *PIK3CA*, *AKT3* or mTOR, who were all reported to have lesions on MRI, became seizure-free after epilepsy surgery (Lee *et al.,* 2012; Poduri *et al.,* 2012; Conti *et al.,* 2015; Jansen *et al.,* 2015; Leventer *et al.,* 2015; Nakashima and Saitsu, 2015). One patient reported some improvement, in another monthly seizures persisted, and the last patient did not become seizure-free, however, outcome was not further specified.

After examination of histology in relation to MRI findings, 19 of the 30 patients (somatic/mosaic or germline combined) had focal cortical dysplasia (FCD) due to mTOR pathway pathogenic variants as a structural substrate of epilepsy, whereas 10 other patients had hemimegalencephaly as the structural cause of their epilepsy. One patient had normal MRI and histology. Epilepsy surgery successfully controlled seizures in eight of the 10 patients with hemimegalencephaly (80%) and in 14 of the 19 patients with FCD (74%). Epilepsy surgery was not successful for the patient with normal MRI and histology.

Overall, epilepsy surgery completely controlled seizures in seven of 12 patients (58%) with epilepsy due to germline mutations in the mTOR

pathway. The success rate was 71% (22 of 30 patients) for germline and somatic mutations combined.

## Epilepsy due to other genetic causes

Eleven articles described a total of 38 patients (all but three were positive for MRI lesions) who had epilepsy in relation to the following other genetic causes: microdeletions, neurofibromatosis type 1, fragile-X syndrome, and mitochondrial mutations (table 4).

Twelve patients who underwent epilepsy surgery have been reported with microdeletions, four of which were identified in 16p13.11 (Catarino *et al.*, 2011; Liu *et al.,* 2012). Nine of 12 patients (75%) became seizure-free after surgery, one patient became seizure-free for seven years after surgery, and the other two patients experienced no improvement.

Twenty-one patients with neurofibromatosis type 1, caused by mutations in *NF1* or microdeletions in 17q11.2 encompassing this gene, underwent epilepsy surgery (Barba *et al.,* 2013; Jang *et al.,* 2013; Ostendorf *et al.,* 2013). These patients had a variety of neurofibromatosis-related epileptogenic lesions, such as HS or low-grade tumours. Epilepsy surgery successfully controlled seizures in 12 of 21 patients (57%) with neurofibromatosis type 1.



Two patients with epilepsy due to Fragile-X syndrome, both with HS, became seizure-free after epilepsy surgery (Wouters *et al.,* 2006; Kenmuir *et al.,* 2015).

Three patients with epilepsy and mitochondrial mutations, who all had HS (detected on MRI), underwent epilepsy surgery; only one became seizure-free (Niehusmann *et al.,* 2011; Azakli *et al.,* 2013).

## MRI-negative patients with genetic epilepsy

A subgroup analysis of all MRI-negative patients with genetic causes of epilepsy yielded a total of 12 patients with mutations (all detected in blood, and not in tissue) in *SCN1A*, *SCN1B*, *CNTNAP2*, *STXBP1*, *DEPDC5*, and *NPRL2*, and microdeletions in 16p13.11, or neurofibromatosis type 1 (table 1 *and* table 5).

Five MRI-negative patients had epilepsy due to mutations in genes related to channelopathies and disorders of synaptic transmission (Strauss *et al.,* 2006; Scheffer *et al.,* 2007; Weckhuysen *et al.,* 2013; Skjei *et al.,* 2015). According to the reports, surgery was considered in these patients based on focal seizure semiology in combination with consistent EEG source localization and results from functional imaging (Scheffer *et al.,* 2007; Skjei *et al.,* 2015), EEG (Weckhuysen *et al.,* 2013) or S-EEG results (Strauss *et al.,* 2006). Surgery did not successfully control seizures in any of these patients, except in one with a mutation in *SCN1B*.

Surgery was performed in four MRI-negative patients who had epilepsy due to mutations in the mTOR pathway genes, DEPDC5 or *NPRL2*, and showed focal abnormalities on EEG, S-EEG or PET (Baulac *et al.,* 2015; Carvill *et al.,* 2015; Weckhuysen *et al.,* 2016). Two MRI-negative patients with *DEPDC5* mutations showed FCD on pathological examination and became seizure-free after surgery, whereas surgery did not successfully control seizures in one patient with a pathology-negative *DEPDC5* mutation and another with a mutation in *NPRL2* and FCD on pathological examination.

Epilepsy surgery did not successfully control seizures in either of the two MRI-negative patients with epilepsy due to a microdeletion in 16p13.11. These patients underwent surgery for clinically presumed HS, although neither had HS on pathological examination (Catarino *et al.,* 2011; Liu *et al.,* 2012). One MRI-negative patient with epilepsy due to neurofibromatosis type 1, with focal abnormalities in the temporal region on EEG and pathology revealing HS in resected tissue, underwent epilepsy surgery and subsequently became seizure-free (Barba *et al.,* 2013).

After histological examination, five of the 12 MRI-negative patients were shown to have features of FCD (type Ia, IIa, or not further specified), one patient had HS, and another had a small epileptogenic hamartoma. Five of the 12 MRI-negative patients (42%) had no abnormalities on histological examination.

Overall, a seizure freedom rate of 33% (4 of 12 patients) was reported in the MRI-negative group; two with a mutation in a mTOR pathway gene, one with a *SCN1B* mutation, and one with an *NF1* mutation. Histological

examination showed a lesion in three of these patients but no abnormality in the patient with the *SCN1B* mutation. One of the five MRI-negative patients without pathological abnormalities became seizure-free after epilepsy surgery, whereas three of seven MRI-negative patients with pathological abnormalities became seizure-free.

## Statistical analyses

Surgery was more successful for patients with mTOR pathway mutations, compared to patients with mutations in genes involved in channelopathies and disorders of synaptic transmission (only patients with germline mutations: 58% versus 14%; Chi-square=5.54; df=1; $p$=0.019; germline and somatic mutations combined: 73% versus 14%; Chi-square=13.42; df=1; $p$<0.001; only patients with MRI-visible lesions: 77% versus 11%; Chi-square= 12.07; df=1; $p$<0.001). The difference in surgery success rate between patients with MRI-visible lesions and MRI-negative patients was at trend level (63% versus 33%, Chi-square=3.679; df=1; $p$=0.055).

| Study | Gene | Mutations | Number of patients | Histology of epileptic focus | MRI findings | Surgery type | Post-surgery seizure outcome | Epilepsy surgery successful | Mean follow-up time in years (range) |
|---|---|---|---|---|---|---|---|---|---|
| **Barba et al., 2014**[15] | SCN1A | c.317C>T; c.2584C>G | 2 | FCD type Ia (1/2); FCD type IIa (1/2) | FCD (2/2) | Right temporal lobectomy and parieto-occipital corticectomy | Engel class IV (1/2), post-operative death (1/2) | 0/2 | Not reported |
| **Skjei et al., 2015** [16] | SCN1A | c.2927del; c.5434T>G; c.4587del; c.5018T>G; c.1661A>G; deletion at exons 17-20 | 6 | Mild MCD (4/6); non-specific findings, possibly due to sampling errors (2/6) | Hippocampal sclerosis (2/6); hemorrhage and encephalomalacia after injury (1/6); small area of abnormal signal (1/6); MRI-negative or non-specific (2/6) | Frontal lobectomy (2/6); temporal lobectomy and frontal gyrus resection (1/6); focal lesionectomy (3/6) | ILAE class 5 (5/6) ILAE class 4 (1/6) | 0/6 | 3.15 (1.5 – 5.9) |
| **Scheffer et al., 2007** [24] | SCN1B | c.363C>G | 2 | Hippocampal sclerosis (1/2); non-specific findings (1/2) | Hippocampal sclerosis (1/2); non-specific (1/2) | Temporal lobectomy | Seizure-free (2/2) | 2/2 | 2.5 (2 – 3) |
| **Strauss et al., 2006** [25] | CNTNAP2 | 3709delG | 3 | Diffuse dysplasia (3/3) | Temporal lobe abnormalities suggestive of MTS (2/3), MRI-negative (1/3) | Temporal lobectomy (1/3), amygdalohippocampectomy (1/3) and limited cortical resection (1/3) | All patients had a recurrence of seizures from 6 to 15 months after surgery | 0/3 | Not reported |
| **Weckhuysen et al., 2013** [23] | STXBP1 | c.1631G>T | 1 | FCD type IA | Normal | Complete occipital disconnection and multiple subpial transections | 95% reduction in seizure frequency (equivalent to Engel class III) | 0/1 | Not reported |

**Table 2: Seizure outcome for patients with epilepsy due to pathogenic variants of genes related to ion channel function and synaptic transmission.** FCD: focal cortical dysplasia; MCD: malformations of cortical development; MRI; magnetic resonance imaging; HS: hippocampal sclerosis; N/A: not applicable; ILAE Class: International League Against Epilepsy classification for seizure outcome following epilepsy surgery.

| Study | Gene | Mutations | Number of patients | Histology of epileptic focus | MRI findings | Surgery type | Post-surgery seizure outcome | Epilepsy surgery successful | Mean follow-up time in years (range) |
|---|---|---|---|---|---|---|---|---|---|
| **Baulac et al., 2015** [26] | DEPDC5 | c.715C>T; c.4841G>A; c.1264C>T; c.1759C>T | 5 | FCD type I (1/5); type IIa (2/5); not conclusive due to fragmented specimen (2/5) | FCD (3/5); MRI-negative (2/5) | Local resection at different sites | Seizure-free (3/5), worthwhile improvement (1/5) no improvement (1/5) | 3/5 | 8 (4 – 13) |
| **Carvill et al., 2015** [27] | DEPDC5 | c.3092C>A; c.842A>T | 2 | Normal (1/2); FCD type IIa (1/2) | Normal (1/2); FCD (1/2) | Temporal corticectomy (1/2); lobectomy followed by functional hemispherectomy (1/2) | No improvement (1/2) and seizure-free for 2.5 years, after which monthly staring spells (1/2) | 0/2 | Not reported |
| **Scerri et al., 2015** [28] | DEPDC5 | c.1663C>T | 2 | FCD type IIa (2/2) | FCD (2/2) | Hemispherectomy (1/2); local resection (1/2) | Seizure-free (2/2) | 2/2 | 13 (10 - 16) |
| **Jansen et al., 2015** [13] | PTEN | p.Tyr68His (germline) | 1 | Hemimegalencephaly and FCD | Hemimegalencephaly | Hemispherectomy | Seizure-free | 1/1 | Not reported |
| | PIK3CA | p.His1047Arg (1/2), p.Glu545Lys as well as p.Thr544Asn (1/2); all mosaic mutations | 2 | Hemimegalencephaly (1/2), FCD type IIa (1/2) | Hemimegalencephaly (1/2); FCD (1/2) | Hemispherectomy (1/2) and occipital resection (1/2) | Seizure-free | 2/2 | Not reported |
| **Conti et al., 2015** [30] | AKT3 | Mosaic trisomy of the 1q21.1-q44 | 1 | FCD type Ib | FCD | Lesionectomy | Seizure-free, except when withdrawing medication | 1/1 | 3 |
| **Lee et al., 2012** [14] | PIK3CA | Somatic p.Glu54Lys mutation | 4 | Hemimegalencephaly and FCD (4/4) | Hemimegalencephaly (4/4) | Hemispherectomy | Seizure-free (4/4) | 4/4 | Not reported |
| | AKT3 | Somatic p.Glu17Lys mutation | 1 | Hemimegalencephaly and FCD | Hemimegalencephaly | Hemispherectomy | Not seizure-free | 0/1 | Not reported |
| | mTOR | Somatic p.Cys1483Tyr mutation | 1 | Hemimegalencephaly and FCD | Hemimegalencephaly | Hemispherectomy | Seizure-free | 1/1 | Not reported |
| **Nakashima et al., 2015** [33] | mTOR | Somatic p.Ser2215Tyr (2/6); p.Ala1459Asp; p.Leu1460Pro; p.Ser2215Phe; p.Leu1460Pro mutations | 6 | FCD (6/6) | FCD (6/6) | Lesionectomy (5/6); amygdalohippocampectomy and temporal lobectomy (1/6) | Seizure-free (5/6); monthly seizures (1/6) | 5/6 | Not reported |
| **Poduri et al., 2012** [31] | AKT3 | Somatic p.E17K activating mutation in AKT3 (1/2) and mosaic trisomy of chromosome 1q involving AKT3 (1/2) | 2 | Hemimegalencephaly and FCD (2/2) | Hemimegalencephaly (2/2) | Hemispherectomy | Seizure-free (1/2); reduction to 1-4 per month (1/2) | 1/2 | 5.5 (5 - 6) |
| **Leventer et al., 2015** [32] | mTOR | Mosaic c.4487T>G mutation | 1 | FCD type IIa | FCD | Hemispherectomy followed by removal of right frontobasal connection | Seizure-free | 1/1 | Not reported |
| **Weckhuysen et al., 2016** [29] | NPRL2 | C68_69delCT | 1 | FCD type Ia | Normal | Fronto-orbital resection | 50% seizure reduction | 0/1 | Not reported |
| | NPRL3 | c.1270C>T | 1 | FCD type IIa | FCD and hippocampal atrophy | Lesionectomy for FCD, twice due to incomplete resection | Rare seizures when medication errors | 1/1 | Not reported |

**Table 3: Seizure outcome after epilepsy surgery in patients with pathogenic variants of mTOR pathway genes.** FCD: focal cortical dysplasia; MRI: magnetic resonance imaging; HME: hemimegalencephaly.

| Study | Genetic cause/syndrome | Number of patients | Histology of epileptic focus | MRI findings | Surgery type | Post-surgery seizure outcome | Epilepsy surgery successful | Mean follow-up time in years (range) |
|---|---|---|---|---|---|---|---|---|
| **Catarino et al., 2011** [35] | Microdeletions in: 16p13.11 (3/10); 17p12 (2/10); 15q11.2 (2/10); 4q32.3 (1/10: 4q35.2 (1/10); 7q31.32-31-33 | 10 | Hippocampal sclerosis (8/10); hamartoma (1/10); non-specific findings (1/10) | Hippocampal sclerosis (7/10), hippocampal atrophy (1/10), hippocampal asymmetry (1/10) and MRI-negative (1/10) | Anterior temporal lobectomy (7/10); amygdalo-hippocampectomy (2/10); neocorticectomy and amygdalectomy (1/10) | ILAE class 1 (8/10); ILAE class 5 (1/10); ILAE 1 for 7 years then ILAE 3 (1/10) | 8/10 | 4.9 (1 – 13) |
| **Liu et al., 2012** [34] | Microdeletions in 16p13.11 | 2 | Hippocampal sclerosis (1/2) and hamartia in white matter of middle temporal gyrus (1/2) | Hippocampal atrophy (1/2) and MRI-negative (1/2) | Hippocampal and temporal cortical resection (2/2) | Seizure free (1/2) and still experiencing frequent seizures (1/2) | 1/2 | Not reported |
| **Barba et al., 2013** [36] | Neurofibromatosis type 1 | 12 | DNET tumour (5/12); hippocampal sclerosis (4/12); polymicrogyria (1/12); mixed pathology (1/12); no pathology (1/12) | Hippocampal sclerosis (3/12); cystic lesion (3/12); cortical dysplasia (2/12); hyperintensities (2/12); polymicrogyria (1/12); MRI-negative (1/12) | Lobectomy (7/12), lesionectomy (1/12), corticectomy (1/12), cortico+lesionectomy (2/12), focal resection+disconnection (1/12) | Engel Ia (8/12), Engel IV (4/12) | 8/12 | 2.6 (1 - 10) |
| **Ostendorf et al., 2013** [37] | Neurofibromatosis type 1 | 8 | Hippocampal sclerosis (2/8); astrocytoma (3/8); glioma (1/8); DNET tumour (1/8); epidermoid (1/8) | MRI and histology not separated in paper | Resection (5/8); lobectomy (3/8) | Engel Ia (3/8), Engel IIb (1/8), Engel IV (4/8) | 3/8 | 7.4 (1 - 33) |
| **Jang et al., 2013** [38] | Neurofibromatosis type 1 | 1 | Hippocampal sclerosis | Hippocampal sclerosis | Anteromesial temporal resection | Seizure-free | 1/1 | Not reported |
| **Wouters et al., 2006** [39] | Fragile-X syndrome: expansion of CGG repeats in FMR1 gene | 1 | Hippocampal sclerosis | Hippocampal sclerosis | Temporal lobectomy | Seizure-free | 1/1 | "Up to this day" |
| **Kenmuir et al., 2015** [40] | Fragile-X syndrome: expansion of CGG repeats in FMR1 gene | 1 | Hippocampal sclerosis | Hippocampal sclerosis | Left anterior temporal lobectomy | Seizure-free | 1/1 | 1 |
| **Niehusmann et al., 2011**[42] | Mitochondrial mutations in ND2 (1/2) and ND4 (1/2) | 2 | Hippocampal sclerosis (1/2); non-specific (1/2) | Hippocampal sclerosis (2/2) | Amygdalohippocampectomy (1/2) and lesionectomy, sparing hippocampus and amygdala (1/2) | Engel class III (1/2); Engel class IIIa (1/2) | 0/2 | Not reported |
| **Azakli et al., 2013**[41] | Mitochondrial mutation in ND1 | 1 | Hippocampal sclerosis | Hippocampal sclerosis | Amygdala hippocampectomy | Seizure-free 3 years after surgery | 1/1 | 3 |

**Table 4: Seizure outcome after epilepsy surgery in patients with epilepsy due to chromosomal abnormalities and genetic syndromes with co-morbid epilepsy.** FCD: focal cortical dysplasia; DNET: dysembryoplastic neuroepithelial tumour; N/A: not applicable; ILAE class: International League Against Epilepsy classification for seizure-outcome following epilepsy surgery.

| Study | Gene | Mutations | Number of patients | Histology of epileptic focus | MRI findings | Surgery type | Post-surgery seizure outcome | Epilepsy surgery successful | Mean follow-up time in years (range) |
|---|---|---|---|---|---|---|---|---|---|
| **Skjei et al., 2015** [16] | SCN1A | c.2927del, c.1661A>G | 2 | No gross abnormalities (2/2); increased cell numbers in 1 patient | Non-specific (2/2) | Frontal lobectomy (1/2) and focal lesionectomy (1/2) | ILAE class 5 (1/2); ILAE class 4 (1/2) | 0/2 | 3.4 (3 - 4) |
| **Scheffer et al., 2007** [24] | SCN1B | C121W | 1 | Non-specific findings | Non-specific scattered and diffuse white matter hyperintensities | Anterior temporal lobectomy | Seizure-free | 1/1 | 2 |
| **Strauss et al., 2006** [25] | CNTNAP 2 | 3709delG | 1 | FCD | No dysplasia | Temporal lobectomy, amygdalohippocampectomy and limited cortical resection | Recurrence of seizures started somewhere between 6 to 15 months after surgery (not further specified) | 0/1 | Not reported |
| **Weckhuysen et al., 2013** [23] | STXBP1 | c.1631G>T | 1 | FCD type Ia | Normal | Complete occipital disconnection and multiple subpial transections | 95% reduction in seizure frequency | 0/1 | Not reported |
| **Baulac et al., 2015** [26] | DEPDC5 | c.484-1G>A (1/2); c.715C>T and c.1264C>T (1/2) | 2 | FCD type I (1/2); FCD type IIa (1/2) | Normal | Local resection | Seizure-free (2/2) | 2/2 | 13 |
| **Carvill et al., 2015** [27] | DEPDC5 | c.3092C>A | 1 | Normal | Normal | Temporal corticectomy | No improvement | 0/1 | Not reported |
| **Weckhuysen et al., 2016** [29] | NPRL2 | C68_69delCT | 1 | FCD type Ia | Normal | Frontoorbital resection | 50% seizure reduction | 0/1 | Not reported |
| **Catarino et al., 2011** [35] | 16p13.11 | chr16:15387380–16198600 microdeletion | 1 | Non-specific findings | Normal | Neocorticectomy and amygdalectomy | ILAE 1 for 7 years then ILAE 3 | 0/1 | 13 |
| **Liu et al., 2012** [34] | 16p13.11 | NDE1-containing microdeletion in resected tissue | 1 | Hamartia in subcortical white matter | Normal | Hippocampal and temporal cortical resection | Still experiencing frequent seizures | 0/1 | Not reported |
| **Barba et al., 2013** [36] | NF-1 | 17q11.2 microdeletion (Neurofibromatosis type 1) | 1 | Hippocampal sclerosis | Normal | Temporal lobectomy | Engel Ia | 1/1 | 2 |

**Table 5: Seizure outcome after epilepsy surgery in the subgroup of MRI-negative patients with genetic epilepsy from the above described studies.** FCD: focal cortical dysplasia; ILAE class: International League Against Epilepsy classification for seizure-outcome following epilepsy surgery.

# Discussion

In this systematic review, we provide an overview of the reported seizure outcomes of patients with different genetic causes of refractory epilepsy who have undergone epilepsy surgery. Not unexpectedly, there was a large difference in success rate of epilepsy surgery between patients with mutations in genes related to channelopathies and disorders of synaptic transmission and those with mutations in the mTOR pathway, even when somatic mutations were excluded for analysis. This difference remains significant when only MRI-positive cases are compared. mTOR pathway genetic variants are thought to increase seizure susceptibility due to abnormal neuronal migration and growth, which leads to (micro) structural epileptogenic malformations of cortical development, such as hemimegalencephaly and FCD (Jansen *et al.*, 2015). Such malformations are thought to arise from a combination of a germline mTOR pathway mutation and a somatic second-hit mutation in the same gene or in a different gene of the mTOR pathway (Poduri *et al.*, 2013; Baulac *et al.*, 2015). This typically results in focal malformations, since the second hit usually only affects part of the brain. It is reasonable to assume that resection of such localised epileptogenic malformations could be a curative treatment for seizures, as reflected by the relatively high surgical success rate of patients with mTOR pathway mutations. It has been estimated that 11% of all focal epilepsies are due to germline mutations in the mTOR genes *DEPDC5*, *NPRL2* and *NPRL3* (Weckhuysen *et al.*, 2016). Considering the associated high success rate of epilepsy surgery, it could be of benefit to routinely screen for such mutations in presurgical evaluation; particularly in MRI-negative, but presumed lesional cases. Finding mTOR pathway mutations would increase the chance of identifying an underlying cryptic malformation of cortical development, and thereby suggest surgical candidacy. The high success rate (83%) of surgery in patients with somatic/mosaic mTOR pathway gene mutations is inherent to the fact that these patients already had established epileptogenic lesions (FCD and hemimegalencephaly); two factors associated with good surgical outcome. Screening for somatic/mosaic mutations in presurgical evaluation is more difficult than for germline mutations. However, investigation of mosaic mutations may be considered in samples of blood, a buccal swab, or sputum using ultra-deep sequencing (Qin *et al.*, 2010).

Epilepsy surgery was almost never successful in patients with epilepsy due to mutations in genes involved in channelopathies and disorders of synaptic transmission. Germline mutations in these genes involved in ion channel function and synaptic transmission are likely to cause widespread aberrant neuronal activity (Helbig *et al.,* 2008), which is rarely confined to a specific part of the brain. It is unlikely that a local resection would be curative to prevent all seizure types. Surgery did not lead to seizure freedom for any reported patient with mutations in *SCN1A*, *CNTNAP2* or *STXBP1* in this series, despite focal semiology for (at least some of their) seizures, and the fact that most of the patients had coincident structural (possibly) epileptogenic lesions. It is likely that these lesions were either not directly related to the genetic cause of epilepsy or that the lesions accounted for only some of the seizures. Possibly, surgery in these patients was not aimed at curing all seizure types, but only targeted seizures originating from a specific structural lesion. However, such a goal was not specified in any of the included articles, nor was the selective outcome for these specific "targeted" seizures. The disappointing overall seizure outcomes of surgery in patients with mutations in this group of genes suggest a relative contraindication for epilepsy surgery, particularly in MRI-negative patients.

Surgery successfully controlled seizures in both patients with mutations in *SCN1B*, one of whom was MRI-negative. In mice, one of the two splice variants of *Scn1b* is known to encode a secreted cell adhesion molecule involved in neuronal pathfinding during embryonic development, and epileptogenic mutations in *Scn1b* result in a functional knockout of this splice variant (Patino *et al.,* 2011). Moreover, *Scn1b* knockout mice exhibit defective neuronal proliferation and migration in the hippocampus, which precedes hyperexcitability (Brackenbury *et al.,* 2013). These findings suggest that *SCN1B* mutations may be associated with structural epileptogenic abnormalities, and focal resection may thus lead to favourable surgery outcome, rather than directly influencing neuronal excitability, as is the case for *SCN1A* mutations (Helbig *et al.,* 2008).

We found large differences in success rate of surgery for epilepsy due to other genetic causes. Epilepsy surgery effectively controlled seizures in most described patients with epilepsy-associated microdeletions. Most of

these patients, however, had HS as an underlying structural epileptogenic substrate, which is generally associated with a favourable surgical outcome. Epilepsy surgery was effective in more than half of the neurofibromatosis type 1 patients. Similar to the situation in patients with pathogenic variants in mTOR pathway genes, *NF1* is thought to affect only those parts of the brain with a second-hit mutation (Poduri *et al.*, 2013), which could explain why resection of these affected parts can be curative. Epilepsy surgery for patients with Fragile-X syndrome or mitochondrial mutations could effectively control seizures, although only a few patients have been described.

Based on a subgroup analysis, we examined whether epilepsy surgery could be effective for MRI-negative patients with genetic epilepsy. Interestingly, MRI-negative patients had a wide range of different genetic causes (table 1), but surgical success rate tended to be higher in cases with MRI showing visible lesions (66%) than in MRI-negative cases (33%), which is in line with previous studies (Téllez-Zenteno *et al.*, 2010). Interestingly, two patients were reported after successful epilepsy surgery for genetic refractory epilepsy, but histological examination of the resected tissue did not reveal any abnormalities. However, we cannot exclude the possibility that subtle abnormalities may have gone undetected due to sampling errors, or that the resection may have removed crucial parts of the epileptogenic non-lesional network. The outcome of these patients suggests that the absence of a detectable lesion on MRI in patients with genetic abnormalities should not in itself be an absolute contraindication for epilepsy surgery.

It remains unclear whether structural lesions are truly absent in MRI-negative patients, or whether their apparent absence is simply based on limitations such as the detection sensitivity threshold of MRI performed or the experience of the radiologist (So and Lee, 2014). In accordance with previous studies (Téllez-Zenteno *et al.*, 2010; So and Lee, 2014), we found that most MRI-negative patients who underwent surgery in this review had histological abnormalities suggestive of MCD in the resected tissue. New MRI methods, higher-field scanning, and post-processing techniques have already shown that it is possible to detect epileptogenic lesions which were not previously visible on conventional MRI scans, improving the

identification of surgical target areas and subsequently yielding higher success rates in patients with genetic refractory epilepsy.

There are a number of limitations to this systematic review. The low number of surgical cases for most genetic causes hampers firm conclusions. Furthermore, there is significant heterogeneity between reported patient characteristics and surgical procedures. The follow-up duration largely varies between studies and is sometimes not reported. Another source of heterogeneity stems from different mutations within the same gene among patients, which could potentially affect surgical outcome. Moreover, differences of expertise in genetic analysis or surgery, accessibility to genetic testing, and indications for epilepsy surgery could relate to lower reporting and different success rates of surgery. Although not explicitly stated in most studies, we assumed that reported mutations were detected in blood, unless specified otherwise. The extent of mosaicism and the effect on the occurrence of a lesion and surgical outcome remains unclear. In addition, publication bias, recall bias, and selection bias due to the scarce number of patients described in the literature cannot be excluded; it is possible that unsuccessful surgery is less likely to be reported.

Surgical candidacy, particularly for MRI-negative patients, is still not easily determined. Some patients are declined surgery because of a presumed non-structural, genetic aetiology. Finding a germline or mosaic mTOR gene mutation could justify continuation of the presurgical diagnostic process. Others, however, are offered resective surgery or invasive monitoring (sEEG) (because of presumed focal structural MRI-negative aetiology), although their epilepsy may have been primarily caused by a genetic and more diffuse aetiology, such as mutations in genes involved in ion channel or neurotransmitter function. Genetic testing is not yet routinely included in most surgical evaluation programmes. Nevertheless, finding specific gene mutations could prove valuable for the process of selecting surgical candidates and counselling patients on expected outcome. Larger and prospective studies are needed to further elucidate the importance of detecting genetic mutations in patients who are considered possible candidates for epilepsy surgery.

## Supplementary data

Summary didactic slides and supplementary tables are available at: https://tinyurl.com/4j3srm5j.

## Acknowledgements and disclosures

We thank the Ming Fund for supporting the "Genetics and targeted treatment in refractory epilepsy" research project. None of the authors have any conflict of interest to declare.

# References

Anderson G.R., Galfin T., Xu W. Candidate autism gene screen identifies critical role for cell-adhesion molecule CASPR2 in dendritic arborization and spine development. *Proc Natl Acad Sci USA*. 2012;109:18120-18125.

Azakli H., Gurses C., Arikan M. Whole mitochondrial DNA variations in hippocampal surgical specimens and blood samples with high-throughput sequencing: a case of mesial temporal lobe epilepsy with hippocampal sclerosis. *Gene*. 2013;529:190-194.

Baldassari S., Licchetta L., Tinuper P., Bisulli F., Pippucci T. GATOR1 complex: the common genetic actor in focal epilepsies. *J Med Genet*. 2016;53:503-510.

Barba C., Jacques T., Kahane P. Epilepsy surgery in Neurofibromatosis Type 1. *Epilepsy Res*. 2013;105:384-395.

Barba C., Parrini E., Coras R. Co-occurring malformations of cortical development and gene mutations. *Epilepsia*. 2014;55:1009-1019. *SCN1A*

Bast T. Outcome after epilepsy surgery in children with MRI-negative non-idiopathic focal epilepsies. *Epileptic Disord*. 2013;15:105-113. 2

Bast T, Kahane P, Jayakar P. Epilepsy surgery in MRI-negative patients. In: A. Arzimanoglou, JH Cross, WD Gaillard *et al*. (editors): Pediatric Epilepsy Surgery. John Libbey Eurotext editions, Paris 2016: 315-328.

Baulac S., Ishida S., Marsan E. Familial focal epilepsy with focal cortical dysplasia due to mutations. *Ann Neurol*. 2015;77:675-683. *DEPDC5*

Bourgeois M., Crimmins D.W., de Oliveira R.S. Surgical treatment of epilepsy in Sturge-Weber syndrome in children. *J Neurosurg*. 2007;106:20-28.

Brackenbury W.J., Yuan Y., O'Malley H.A., Parent J.M., Isom L.L. Abnormal neuronal patterning occurs during early postnatal brain development of Scn1b-null mice and precedes hyperexcitability. *Proc Natl Acad Sci USA*. 2013;110:1089-1094.

Carvill G.L., Crompton D.E., Regan B.M. Epileptic spasms are a feature of DEPDC5 mTORopathy. *Neurol Genet*. 2015;1:e17.

Catarino C.B., Kasperaviciute D., Thom M. Genomic microdeletions associated with epilepsy: not a contraindication to resective surgery. *Epilepsia*. 2011;52:1388-1392.

Conti V., Pantaleo M., Barba C. Focal dysplasia of the cerebral cortex and infantile spasms associated with somatic 1q21.1-q44 duplication including the AKT3 gene. *Clin Genet*. 2015;88:241-247.

de Flon P., Kumlien E., Reuterwall C., Mattsson P. Empirical evidence of underutilization of referrals for epilepsy surgery evaluation. *Eur J Neurol*. 2010;17:619-625.

Dewar S.R., Pieters H.C. Perceptions of epilepsy surgery: a systematic review and an explanatory model of decision-making. *Epilepsy Behav*. 2015;44:171-178.

Helbig I., Scheffer I.E., Mulley J.C., Berkovic S.F. Navigating the channels and beyond: unravelling the genetics of the epilepsies. *Lancet Neurol*. 2008;7:231-245.

Hildebrand M.S., Dahl H-HM, Damiano J.A. Recent advances in the molecular genetics of epilepsy. *J Med Genet*. 2013;50:271-279.

Jang H.M., Park H.R., Mun J.K. Surgical treatment of mesial temporal lobe epilepsy in a patient with neurofibromatosis type 1. *J Epilepsy Res*. 2013;3:35-38.

Jansen L.A., Mirzaa G.M., Ishak G.E. PI3K/AKT pathway mutations cause a spectrum of brain malformations from megalencephaly to focal cortical dysplasia. *Brain*. 2015;138:1613-1628.

Kenmuir C., Richardson M., Ghearing G. Surgical treatment for medically refractory focal epilepsy in a patient with fragile X syndrome. *Brain Dev*. 2015;37:916-918.

Lee J.H., Huynh M., Silhavy J.L. somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nat Genet*. 2012;44:941-945. *De novo*

Leventer R.J., Scerri T., Marsh A.P.L. Hemispheric cortical dysplasia secondary to a mosaic somatic mutation in MTOR. *Neurology*. 2015;84:2029-2032.

Liu J.Y.W., Kasperaviciute D., Martinian L., Thom M., Sisodiya S.M. Neuropathology of 16p13.11 deletion in epilepsy. *PLoS One*. 2012;7:e34813.

Nakashima M., Saitsu H. Somatic mutations in the gene cause focal cortical dysplasia type IIb. *Ann Neurol*. 2015;78:375-386. *MTOR*

Niehusmann P., Surges R., von Wrede R.D. Mitochondrial dysfunction due to Leber's hereditary optic neuropathy as a cause of visual loss during assessment for epilepsy surgery. *Epilepsy Behav*. 2011;20:38-43.

Ostendorf A.P., Gutmann D.H., Weisenberg J.L.Z. Epilepsy in individuals with neurofibromatosis type 1. *Epilepsia*. 2013;54:1810-1814.

Patino G.A., Brackenbury W.J., Bao Y. Voltage-gated Na+ channel beta1B: a secreted cell adhesion molecule involved in human epilepsy. *J Neurosci*. 2011;31:14577-14591.

Poduri A., Evrony G.D., Cai X. Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron*. 2012;74:41-48.

Poduri A., Evrony G.D., Cai X., Walsh C.A. Somatic mutation, genomic variation, and neurological disease. *Science*. 2013;341:1237758.

Prabowo A.S., Iyer A.M., Veersema T.J. BRAF V600E mutation is associated with mTOR signaling activation in glioneuronal tumors. *Brain Pathol*. 2014;24:52-66.

Qin W., Kozlowski P., Taillon B.E. Ultra deep sequencing detects a low rate of mosaic mutations in tuberous sclerosis complex. *Human Genet*. 2010;127:573-582.

Scerri T., Riseley J.R., Gillies G. Familial cortical dysplasia type IIA caused by a germline mutation in . *Ann Clin Transl Neurol*. 2015;2:575-580. *DEPDC5*

Scheffer I.E., Harkin L.A., Grinton B.E. Temporal lobe epilepsy and GEFS+ phenotypes associated with mutations. *Brain*. 2007;130:100-109. *SCN1B*

Skjei K.L., Church E.W., Harding B.N. Clinical and histopathological outcomes in patients with mutations undergoing surgery for epilepsy. *J Neurosurg Pediatr*. 2015;16:1-7. *SCN1A*

So E.L., Lee R.W. Epilepsy surgery in MRI-negative epilepsies. *Curr Opin Neurol*. 2014;27:206-212.

Strauss K.A., Puffenberger E.G., Huentelman M.J. Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N Engl J Med*. 2006;354:1370-1377.

Téllez-Zenteno J.F., Ronquillo L.H., Moien-Afshari F., Wiebe S. Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis. *Epilepsy Res*. 2010;89:310-318.

Uijl S.G., Leijten F.S., Moons K.G. Epilepsy surgery can help many more adult patients with intractable seizures. *Epilepsy Res*. 2012;101:210-216.

Varea O., Martin-de-Saavedra M.D., Kopeikina K.J. Synaptic abnormalities and cytoplasmic glutamate receptor aggregates in contactin associated protein-like 2/Caspr2 knockout neurons. *Proc Natl Acad Sci USA*. 2015;112:6176-6181.

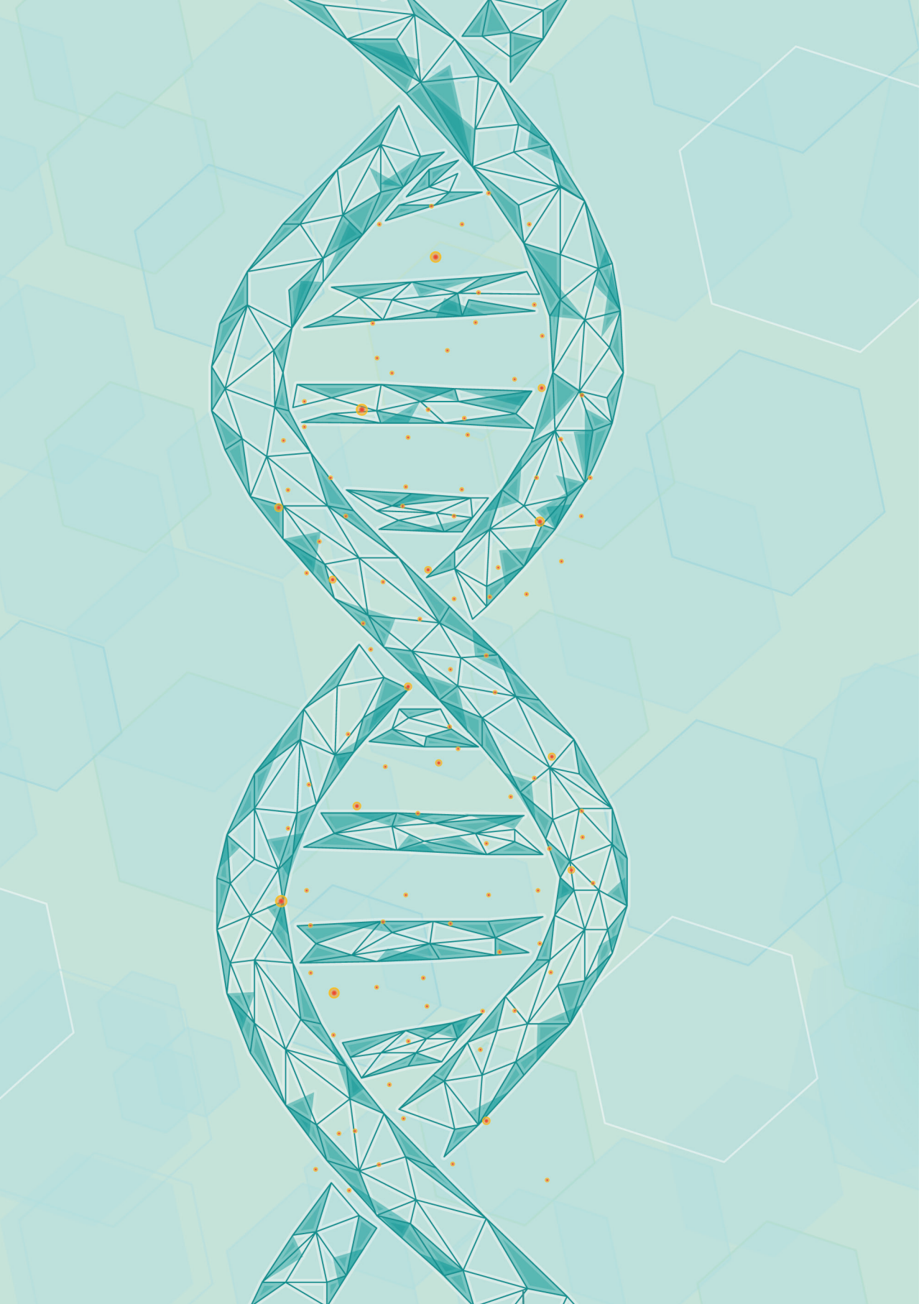
Weckhuysen S., Holmgren P., Hendrickx R. Reduction of seizure frequency after epilepsy surgery in a patient with STXBP1 encephalopathy and clinical description of six novel mutation carriers. *Epilepsia*. 2013;54:e74-80.

Weckhuysen S., Marsan E., Lambrecq V. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia*. 2016;57:994-1003.

West S., Nolan S.J., Cotton J. Surgery for epilepsy. *Cochrane Database Syst Rev*. 2015;7:CD010541.

Wouters H., Fonteyne A., Lagae L., Stiers P. Specific memory impairment in a multiple disabled male with fragile X syndrome and temporal lobe epilepsy. *Dev Med Child Neurol*. 2006;48:378-382.

Zhang K., Hu W.H., Zhang C. Predictors of seizure freedom after surgical management of tuberous sclerosis complex: a systematic review and meta-analysis. *Epilepsy Res*. 2013;105:377-383.



10

## CHAPTER 10

# REFRACTORY JUVENILE MYOCLONIC EPILEPSY: A META-ANALYSIS OF PREVALENCE AND RISK FACTORS

R. Stevelink, B.P.C. Koeleman, J.W. Sander, F.E. Jansen, K.P.J. Braun

*European Journal of Neurology*. 2019; Jun;26(6):856-864.

# Abstract

**Background and purpose**

Juvenile myoclonic epilepsy (JME) is a common epilepsy syndrome for which treatment response is generally assumed to be good. We aimed to determine the prevalence and prognostic risk factors for refractoriness of JME.

**Methods**

We systematically searched PubMed and EMBASE and included 43 eligible studies, reporting seizure outcome after antiepileptic drug (AED) treatment in JME cohorts. We defined refractory JME as persistence of any seizure despite AED treatment and performed a random-effects meta-analysis to assess the prevalence of refractory JME and of seizure recurrence after AED withdrawal in individuals with well-controlled seizures. Studies reporting potential prognostic risk factors in relation to seizure outcome were included for subsequent meta-analysis of risk factors for refractoriness.

**Results**

Overall, 35% (95% confidence interval, 29–41%) of individuals (n = 3311) were refractory. There was marked heterogeneity between studies. Seizures recurred in 78% (95% confidence interval, 52–94%) of individuals who attempted to withdraw from treatment after a period of seizure freedom (n = 246). Seizure outcome by publication year suggested that prognosis did not improve over time. Meta-analysis suggested six variables as prognostic factors for refractoriness, i.e. having three seizure types, absence seizures, psychiatric comorbidities, earlier age at seizure onset, history of childhood absence epilepsy and praxis-induced seizures.

**Conclusion**

One-third of people with JME were refractory, which is a higher prevalence than expected. Risk factors were identified and can be used to guide treatment and counselling of people with JME.

# Introduction

Juvenile myoclonic epilepsy (JME) is the most common form of genetic generalized epilepsy, affecting 5–10% of all people with epilepsy, with a prevalence of 0.1–0.2/100 000[1]. JME typically manifests during adolescence and is characterized by arrhythmic myoclonic seizures, particularly occurring on awakening, and electroencephalography that shows generalized spike and polyspike waves[2]. Although not required for diagnosis, people with JME often also experience generalized tonic–clonic seizures and, less often, absence seizures[2]. According to its definition 'response to appropriate drugs is good'[2]. This could lead to optimistic counselling by physicians. Seizures, however, continue despite adequate treatment with antiepileptic drugs (AEDs) in a proportion of patients and this impacts on quality of life[3,4]. Once an individual becomes seizure-free on AEDs, it is usually recommended to continue life-long therapy, given the high risk of relapse following drug withdrawal[5, 6]. Some studies have suggested that a subset of individuals remains seizure-free after drug withdrawal[7, 8]. It is important to establish how often individuals are refractory and how frequently AEDs can be safely withdrawn to allow reliable prognostic counselling.

Several studies have explored risk factors for refractory JME but individual studies are limited by relatively small sample sizes and there are inconsistencies between studies. Prediction of refractoriness is of value for individualized management, e.g. by considering higher drug doses, polytherapy, experimental AEDs or non-pharmacological treatment options earlier in those at risk[9–12].



We aimed to provide a systematic overview of refractory JME and its prognostic risk factors. By meta-analysing available studies, we estimated the proportion of refractory JME and, at the other end of the spectrum, the proportion of individuals remaining seizure-free after drug withdrawal. Lastly, we assessed which clinical variables may predict refractory JME.

# Methods

## Search strategy and study selection

Procedures were consistent with PRISMA guidelines[13]. A literature search in PubMed and EMBASE identified articles describing treatment outcome in people with JME (see Tables S1 and S2 for search terms). We did not adopt a registered pre-specified protocol.

We included all retrospective and prospective studies reporting seizure outcome after AED treatment in observational cohorts of individuals with a diagnosis of JME, regardless of the diagnostic criteria used by the study (see Table S3 for an overview), which may vary[14]. We excluded articles that specifically recruited refractory individuals or those in remission. Drug-trial reports were not included as they could be biased towards individuals with a refractory condition. We contacted authors of articles describing multiple generalized epilepsy syndromes to provide stratified data of individuals with JME, if not available in the publication. We only included articles describing seizure freedom from all seizure types and excluded those with ambiguous definitions (e.g. 'good outcome') without specifying seizure freedom. When the same cohort was included in multiple reports, we included the most recent report, except in cases where an older article provided data on potential risk factors of refractory JME. Articles in English, Dutch and German were included.

Definitions of seizure freedom and refractory JME varied between articles, primarily regarding the length of the seizure-free follow-up period. Only two articles used the definition of drug-resistant epilepsy proposed by the International League Against Epilepsy in 2010[15]. We defined refractory JME as persistence of any seizure (i.e. myoclonic, absence or generalized tonic-clonic seizures) despite AED treatment, regardless of the length of the seizure-free follow-up period. We assessed 1-year seizure freedom when multiple time points were described within the same study. Where possible, individual cases of 'pseudo-refractory' individuals (i.e. those who had seizures due to non-compliance, inadequate treatment or other factors not related to therapy) were excluded.

Studies reporting potential prognostic risk factors stratified by seizure outcome were included for subsequent meta-analysis of risk factors for refractoriness.

All search results were reviewed based on title and abstract, and the full text was reviewed in potentially eligible articles. Reference lists were checked for additional eligible articles.

## Data extraction

Study selection and data extraction were performed by R.S. A standardized data extraction form was created that contained the number of individuals who were seizure-free and those who were refractory, seizure outcome after drug withdrawal, mean follow-up duration, country, prospective or retrospective design, type of AED used and definition of seizure freedom.

Data of prognostic risk factors from articles reporting clinical variables stratified by seizure outcome were also extracted. To reduce publication bias, raw data of potential risk factors were extracted from all articles, regardless of whether the variable was tested for association with seizure outcome. We analysed only potential risk factors that were reported in at least two articles, regardless of whether they were significantly associated with outcome.



## Statistical analyses

A random-effects meta-analysis was performed using the R package Metafor (v2.0-0) to assess the prevalence of refractoriness. The $I^2$ statistic was assessed as a measure to quantify heterogeneity, where values between 50% and 75% are considered to represent moderate heterogeneity and those >75% represent high heterogeneity[16]. We used a random-effects model to account for heterogeneity between studies[17]. Secondary analyses stratified by definition of refractory JME and by study design (prospective or retrospective) were performed to assess whether this increased homogeneity. Differences by publication year and differences between 1-, 2- and 5-year seizure freedom were assessed with a mixed-effects meta-

regression, using Metafor. A random-effects meta-analysis was performed using Metafor to assess the prevalence of individuals who remained seizure-free after AED withdrawal.

Random-effects meta-analyses of potential risk factors were performed using Review Manager (v5.3) for all potential risk factors reported in at least two articles. We assessed the odds ratio as outcome measure for dichotomous variables and the mean difference for continuous variables.

### Quality and bias assessment

The Newcastle–Ottawa quality assessment scale for cohort studies was used to assess the methodological quality of all studies included in the meta-analysis of risk factors[18]. This scale is used to assess three major components, i.e. cohort selection, comparability and assessment of outcome, and ranges from 0 to 9, where studies are considered to have a high quality when scoring ≥5 and a low quality when scoring <5.

Funnel plots were generated as a measure to assess potential publication bias and were visually inspected for asymmetry[19]. Considering the small number of studies included per risk factor, we did not perform statistical tests for asymmetry of the funnel plot, as it is only recommended when including >10 studies per analysis[19].

## Results

The literature search was last performed on 1 March 2018 and yielded 1362 articles (see Fig. 1 for flowchart). After removing duplicates and applying inclusion and exclusion criteria, 43 articles were included, describing treatment outcomes for a total of 3311 subjects (Table S4).

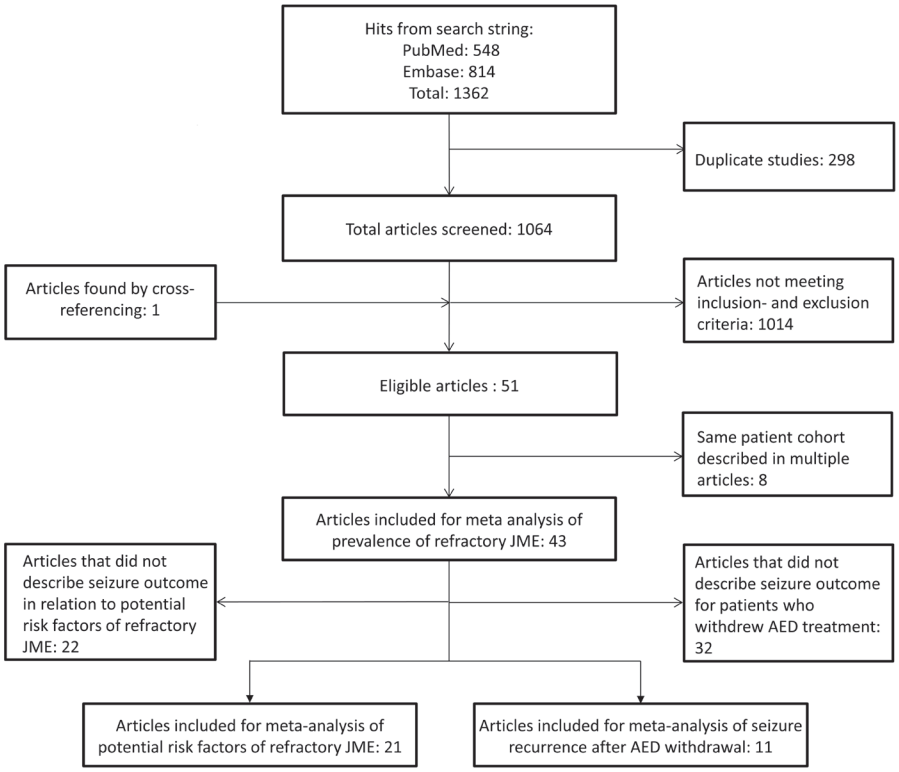



**Figure 1: Flowchart of search strategy and study selection. AED,** antiepileptic drug; JME, juvenile myoclonic epilepsy.



## Prevalence of refractory juvenile myoclonic epilepsy

Meta-analysis showed that 35% [95% confidence interval (CI), 29–41%] of individuals with JME were refractory to treatment (Fig. 2). The proportion of refractory subjects varied between 7% and 75%, and heterogeneity between studies was high (I2 = 91%). As the definition of seizure freedom varied between studies, we also performed analyses stratified by definition, which made little difference to the estimate of refractory JME or the amount of heterogeneity (Fig. 3). A meta-regression analysis showed no significant difference between 1-, 2- and 5-year seizure freedom (P = 0.41). The proportion of refractory patients was comparable between prospective (36%; 95% CI, 18–56%) and retrospective (35%; 95% CI, 29–42%) studies.

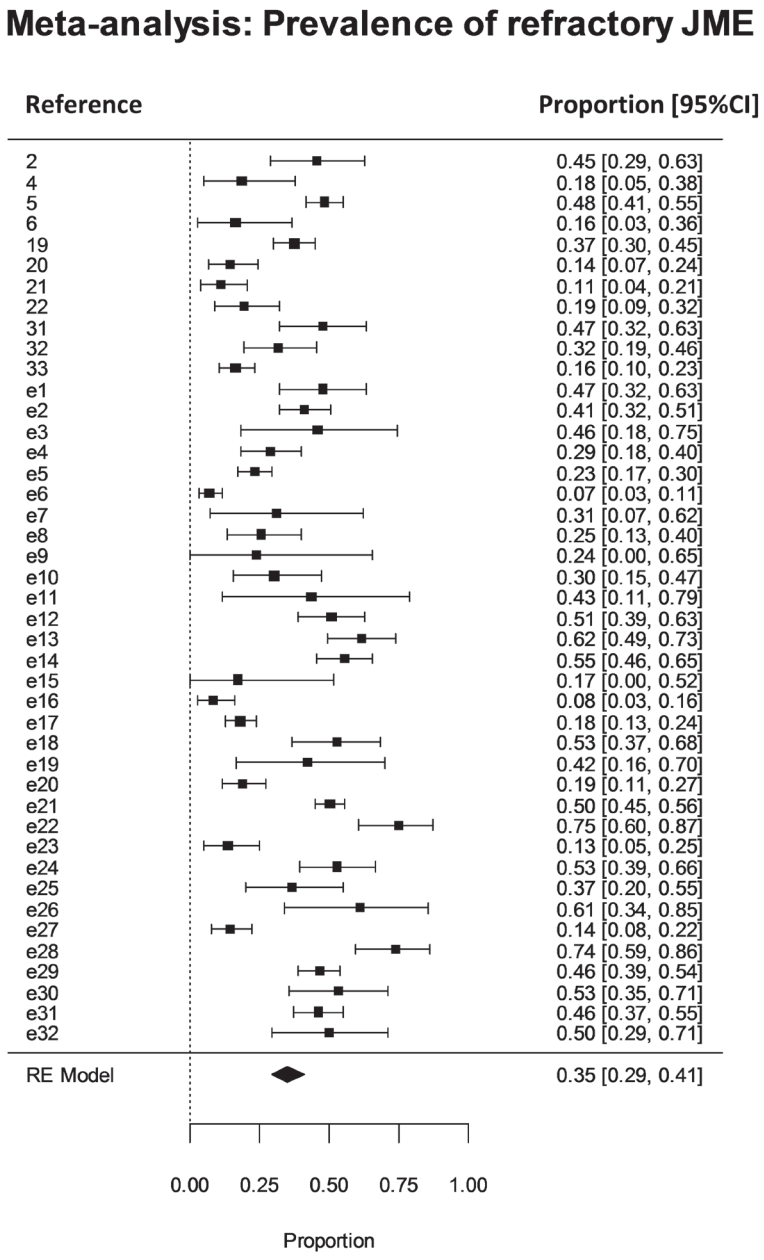



**Figure 2: Meta-analysis of the prevalence of refractory juvenile myoclonic epilepsy (JME).** The proportion of subjects who were refractory is displayed on the x-axis. A total of 43 studies describing seizure outcome in 3311 individuals with JME were included. CI, confidence interval; RE, random-effects. References denoted as 'e' are available in the Supporting Information.

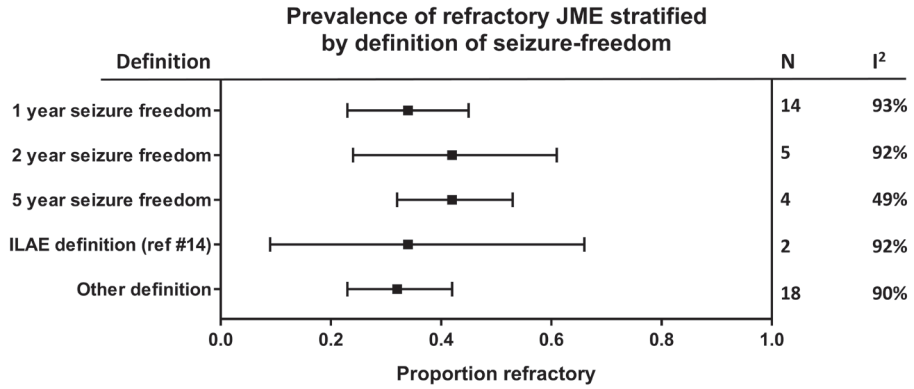



**Figure 3: Meta-analyses of the prevalence of refractory juvenile myoclonic epilepsy stratified by definition of seizure freedom.** ILAE, International League Against Epilepsy; N, number of studies; I2, heterogeneity.

We next assessed whether the proportion of seizure-free individuals has changed over time (Fig. 4). A meta-regression analysis showed no significant association between publication year and percentage of refractoriness (mixed-effects meta-regression: P = 0.61).

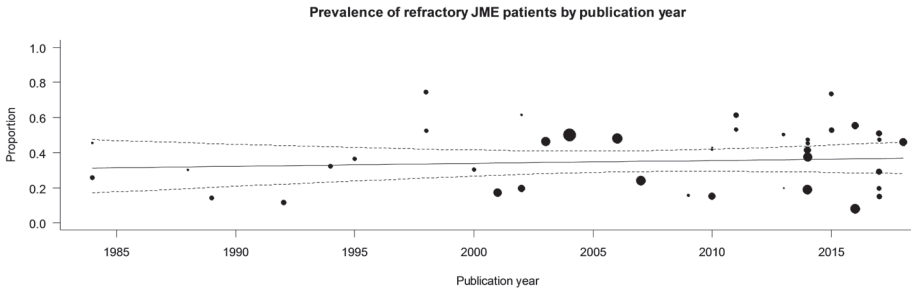



**Figure 4: Meta-regression of refractory juvenile myoclonic epilepsy by publication year. The proportion of refractory subjects per study is plotted by publication year.** Each study is represented by a circle whose size is proportional to the sample size. A meta-regression trend line with 95% confidence interval (dotted lines) is plotted as a solid line.

## Seizure recurrence after antiepileptic drug withdrawal

A total of 11 articles described a subset of 246 subjects who attempted AED withdrawal. Some studies had specific criteria for subjects to withdraw (e.g. at least 3-year seizure freedom), but most did not. Meta-analysis

showed that seizures recurred in 78% (95% CI, 58–94%) of subjects after withdrawal (Fig. 5), although estimates varied widely and heterogeneity was high (I2 = 84%).

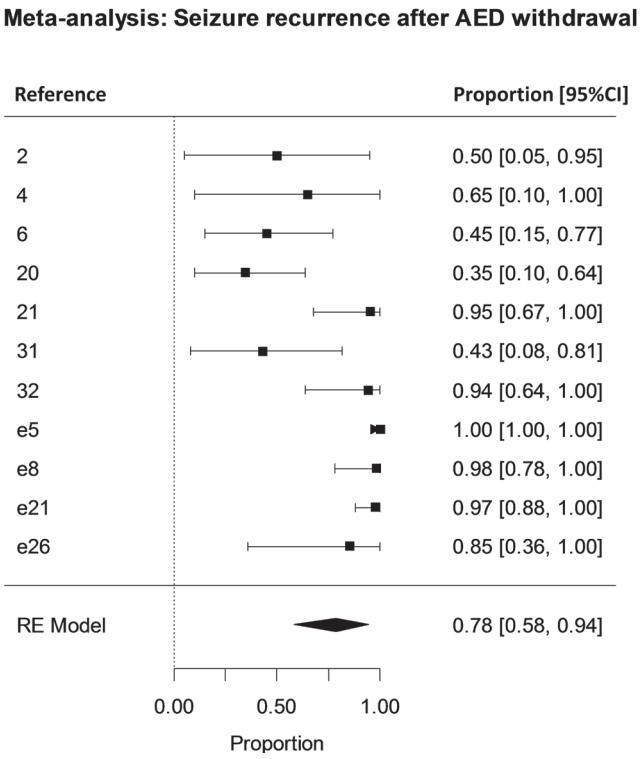



**Figure 5: Meta-analysis of seizure recurrence after antiepileptic drug (AED) withdrawal.** The proportion of well-controlled subjects who experienced recurrence of seizures after AED withdrawal is displayed on the x-axis. A total of 11 studies describing 246 subjects were included. CI, confidence interval; RE, random-effects. References denoted as 'e' are available in the Supporting Information.

## Risk factors for refractory juvenile myoclonic epilepsy

A total of 21 studies reported seizure outcome in relation to potential risk factors for refractory JME. Univariate meta-analyses were performed for 10 risk factors (Table 1; see Figs S1–S10 for forest plots). Having three seizure types, absence seizures, psychiatric comorbidities, a history of childhood absence epilepsy (CAE) progressing to JME, praxis-induced seizures (seizures and epileptiform electroencephalographic discharges precipitated

by complex, cognition-guided tasks, such as playing chess, writing or drawing) and early age at epilepsy onset were each significant risk factors for refractory JME. Heterogeneity between studies was mild to moderate. Scores on the Newcastle–Ottawa quality assessment scale (Table S5) ranged between 2 and 7 (mean 4.1) [13 studies were assessed as low (score ≤4) and 8 as high (score ≥5) quality]. Funnel plots, inspected as a measure of publication bias, did not show asymmetry (Figs S1–S10).

| Risk factor | Number of studies | Number of subjects | Test statistic (95% CI) | P-value | Heterogeneity (I2) |
|---|---|---|---|---|---|
| Three seizure types (myoclonic+GTCS+absences) | 11 | 864 | **OR: 2.97 (1.87, 4.71)** | **<0.00001** | 19% |
| Absence seizures | 13 | 961 | **OR: 2.81 (1.77, 4.45)** | **<0.0001** | 42% |
| Psychiatric comorbidities | 8 | 802 | **OR: 3.78 (2.46, 5.81)** | **<0.00001** | 9% |
| Female gender | 10 | 855 | OR: 1.19 (0.85, 1.66) | 0.32 | 0% |
| Epileptiform asymmetries on EEG | 7 | 622 | OR: 1.66 (0.71, 3.92) | 0.24 | 54% |
| Photoparoxysmal response | 5 | 395 | OR: 0.89 (0.49, 1.62) | 0.70 | 0% |
| Family history of epilepsy | 9 | 782 | OR: 1.03 (0.72, 1.49) | 0.86 | 0% |
| History of childhood absence epilepsy progressing to JME | 4 | 353 | **OR: 4.55 (1.38, 15.01)** | **0.01** | 56% |
| Praxis induced seizures | 2 | 110 | **OR: 3.73 (1.44, 9.68)** | **0.007** | 0% |
| Early age at epilepsy onset | 8 | 517 | **MD: -1.60 (-2.81, -0.40)** | **0.009** | 47% |

**Table 1: Risk factors for refractory JME, assessed with random-effects meta-analysis.** CI, confidence interval; EEG, electroencephalography; GTCS, generalized tonic-clonic seizures; MD, mean difference; OR, odds ratio. Significant associations, defined as a meta-analysis P-value <0.05, are highlighted in bold.

# Discussion

One-third of the described subjects with JME were refractory (Fig. 2). The estimates of refractoriness were comparable when assessing 1-, 2- and 5-year seizure freedom (Fig. 3), suggesting that people who are seizure-free for at least 1 year are likely to remain so. This is consistent with studies

that reported 1- and 2-year or 1- and 5-year seizure freedom in the same subjects, which showed minor differences between outcomes at different follow-up intervals[20, 21].

We found no evidence for a decrease in the proportion of refractory JME over the last decades. Valproate, marketed as an AED since 1967, is still considered the most effective drug for people with JME[9, 22, 23]. Thus, there is still much room for improvement.

In contrast to the International League Against Epilepsy definition (1989) of JME, describing the treatment response to 'appropriate drugs' as 'good', our results suggest that the proportion of refractoriness is not much different from the overall proportion of refractoriness in people with epilepsy, which is estimated between 16% and 37%[24-26]. Physicians should be careful when counselling people with JME that their prognosis is particularly good. It is possible, however, that we overestimated refractoriness in JME. Individuals in the included studies were mainly treated at tertiary centres and are likely to have more severe or difficult-to-treat epilepsy than those at secondary centres. Conversely, it has been shown that seizure control improves after referral to tertiary care[27]. It is also possible that some were misdiagnosed, as other conditions may mimic JME[28]. There is also the possibility of selection bias and selective loss to follow-up of people with a more benign course, who might be less inclined to return to the clinic or agree to inclusion in a study. Our estimate, however, could be an underestimation of refractoriness of myoclonic seizures, which are difficult to objectify and can be under-reported. Another limitation is that study selection and data extraction were performed by a single author. Statistical heterogeneity between studies was substantial, but definition of seizure freedom, publication year or retrospective versus prospective study design did not seem to play a major role in heterogeneity. Other potential causes of heterogeneity could not be assessed, such as ethnic origins, different treatment regimens and different diagnostic criteria. Determining seizure freedom is subjective and a recent study established that inter-observer variability (using the same criteria and the same individual records) was relatively high, with kappa values ranging between 0.56 and 0.77 . It is likely that intra-observer variability would be even higher when the same individual records are not used. Thus,

intra-observer variability is likely to have played a role in heterogeneity between studies.

About one-fifth of subjects are reported to remain seizure-free after treatment withdrawal (Fig. 5), which is substantially less than the overall estimate of two-thirds for all types of epilepsy[30, 31]. Estimates between studies, however, varied widely. A potential cause of heterogeneity is age at withdrawal and therefore duration of seizure freedom, as these variables are predictors of seizure recurrence in the general epilepsy population[30] and JME has been shown to subside with age[32]. Age at AED withdrawal was rarely reported, but the three studies reporting a good prognosis mostly included people over 40 years of age[7, 21, 33], whereas the two studies reporting that all subjects had seizure recurrence included mainly people in their twenties[22, 34]. It is possible that the actual proportion of seizure freedom after AED withdrawal is higher for older subjects. Insufficient information about individuals who attempted AED withdrawal was available to allow identification of potential prognostic factors. Future studies are needed to evaluate which subjects are most likely to remain seizure-free after treatment withdrawal.

Our meta-analyses revealed six significant risk factors for refractoriness, but did not provide evidence for the other four clinical variables to be significantly associated (Table 1). It is likely that these variables are inter-related. For example, a history of CAE relates to having absence seizures and to an earlier age at epilepsy onset[6], and most people with JME who have absence seizures had three seizure types[35].



Cause and effect cannot be established due to the cross-sectional nature of the studies. We cannot rule out that psychiatric comorbidities are due to AED side-effects or to having prolonged refractory seizures, rather than being the cause. It is also possible that people with psychiatric comorbidities are less adherent to treatment rather than being non-responsive to AEDs.

It remains uncertain whether the risk factors for refractory JME represent a lack of response to treatment or a higher disease burden. People with early disease onset, multiple seizure types and psychiatric comorbidities may have more severe brain disease, which makes it more difficult to control all seizure types. Conversely, someone with only occasional seizures can be well

controlled even when the medication is only mildly effective. It has also been suggested that people with CAE progressing into JME represent a distinct clinical entity, with a different inheritance pattern and seizure outcome[6]. They rarely become completely free of all seizures. Most described individuals, however, do become free of myoclonic seizures and generalized tonic–clonic seizures, with only absences persisting. This suggests the possibility that different seizure types respond differently to treatment. A genetic study comparing drug-responsive individuals with those who are refractory could unravel a distinct genetic basis of treatment response, higher genetic overlap with CAE or higher polygenic burden of JME-associated risk alleles.

Further studies using individual data are required to assess which variables are independent predictors of refractory JME to allow for an individualized prediction of seizure outcome to be used to guide treatment.

## Acknowledgements

We are grateful to the Ming Fund for supporting this project. We thank Dr Pierre Genton for valuable discussions and advice on the analyses and interpretations. We also thank Dr Christoph Beier, Dr Bernd Vorderwülbecke, Dr Philine Senf, Dr Giorgi Japaridze and Dr Katie Holland for providing stratified data.

## Disclosure of conflicts of interest

The authors declare no financial or other conflicts of interest. J.W.S. reports grants and personal fees from Eisai, grants and personal fees from UCB, grants from WHO, grants from NEF, personal fees from Eisai, and grants and personal fees from UCB, outside the submitted work. His current position is endowed by the Epilepsy Society and he is a member of the Editorial Board of the *Lancet Neurology* and receives research support from the Marvin Weil Epilepsy Research Fund.

## Supporting information

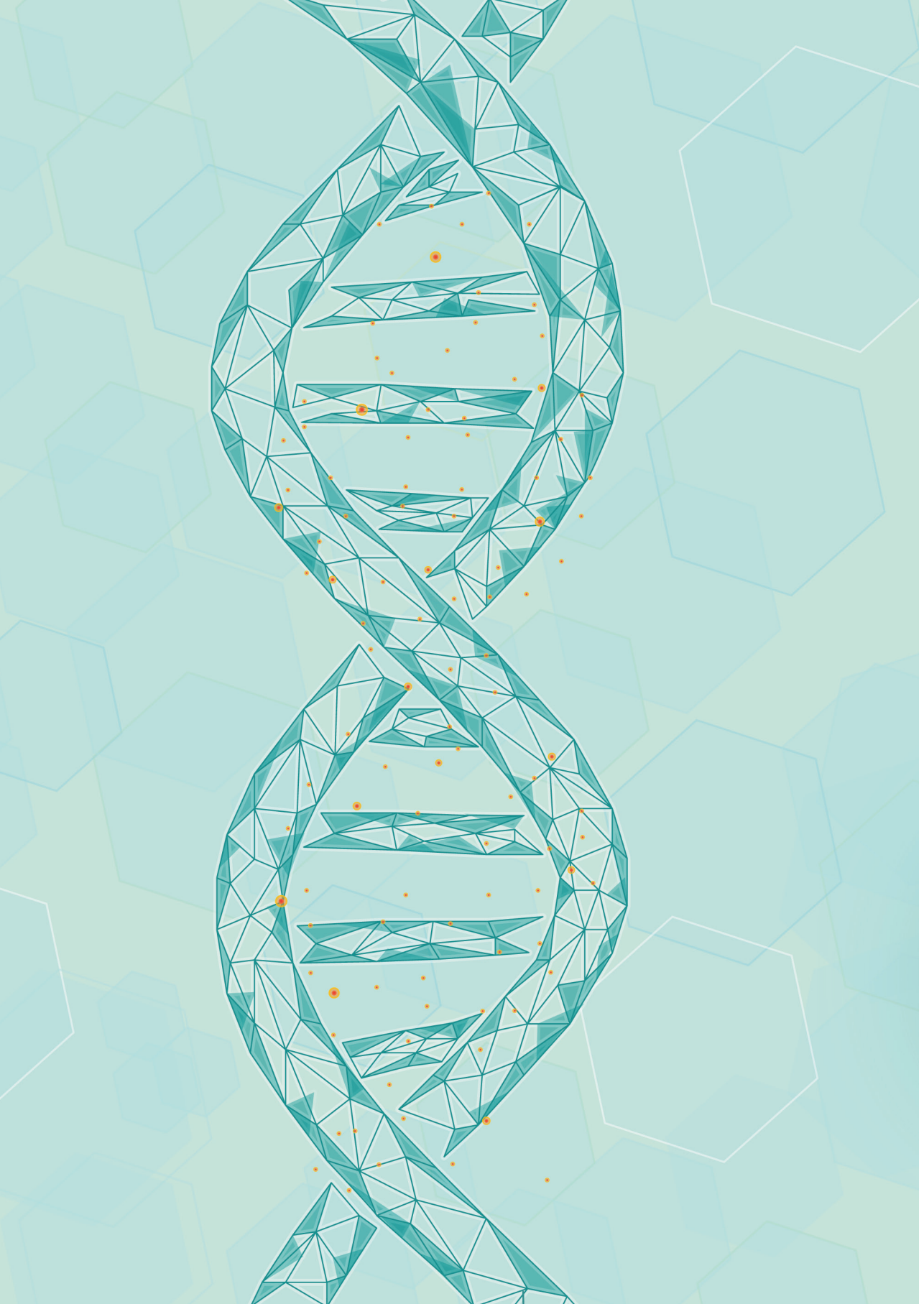
Additional Supporting Information may be found online at: https://tinyurl.com/4j3srm5j.

# References


1. Camfield CS, Striano P, Camfield PR. Epidemiology of juvenile myoclonic epilepsy. *Epilepsy Behav* 2013; **28**(Suppl. 1): S15– S17.
2. Commission on Classification and Terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia* 1989; **30**: 389– 399.
3. Schneider-von Podewils F, Gasse C, Geithner J, *et al*. Clinical predictors of the long-term social outcome and quality of life in juvenile myoclonic epilepsy: 20-65 years of follow-up. *Epilepsia* 2014; **55**: 322– 330.
4. Leidy NK, Elixhauser A, Vickrey B, Means E, Willian MK. Seizure frequency and the health-related quality of life of adults with epilepsy. *Neurology* 1999; **53**: 162– 166.
5. Calleja S, Salas-Puig J, Ribacoba R, Lahoz CH. Evolution of juvenile myoclonic epilepsy treated from the outset with sodium valproate. *Seizure* 2001; **10**: 424–427.
6. Martínez-Juárez IE, Alonso MEME, Medina MT, *et al*. Juvenile myoclonic epilepsy subsyndromes: family studies and long-term follow-up. *Brain* 2006; **129**: 1269–1280.
7. Camfield CS, Camfield PR. Juvenile myoclonic epilepsy 25 years after seizure onset: a population-based study. *Neurology* 2009; **73**: 1041– 1045.
8. Senf P, Schmitz B, Holtkamp M, *et al*. Prognosis of juvenile myoclonic epilepsy 45 years after onset: seizure outcome and predictors. *Neurology* 2013; **81**: 2128–2133.
9. Nicolson A, Marson AG. When the first antiepileptic drug fails in a patient with juvenile myoclonic epilepsy. *Pract Neurol* 2010; **10**: 208– 218.
10. Mantoan L, Walker M, Mantoan L, *et al*. Treatment options in juvenile myoclonic epilepsy. *Curr Treat Options Neurol* 2011; **13**: 355– 370.
11. Kossoff EH, Henry BJ, Cervenka MC. Efficacy of dietary therapy for juvenile myoclonic epilepsy. *Epilepsy Behav* 2013; **26**: 162– 164.
12. Jenssen S, Sperling MR, Tracy JI, *et al*. Corpus callosotomy in refractory idiopathic generalized epilepsy. *Seizure* 2006; **15**: 621– 629.
13. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.
14. Kasteleijn-Nolst Trenité DGA, Schmitz B, Janz D, *et al*. Consensus on diagnosis and management of JME: from founder's observations to current trends. *Epilepsy Behav* 2013; **28**: S87– S90.
15. Kwan P, Arzimanoglou A, Berg AT, *et al*. Definition of drug resistant epilepsy: consensus proposal by the *ad hoc* Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia* 2010; **51**: 1069– 1077.
16. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539– 1558.
17. Hunter JE, Schmidt FL. Fixed effects vs. random effects meta-analysis models: implications for cumulative research knowledge. *Int J Sel Assess* 2000; **8**: 275– 292.

18. Wells G, Shea B, O'Connell D, *et al*. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 15/03/2018).
19. Sterne JAC, Sutton AJ, Ioannidis JPA, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002.
20. Höfler J, Unterberger I, Dobesberger J, Kuchukhidze G, Walser G, Trinka E. Seizure outcome in 175 patients with juvenile myoclonic epilepsy - A long-term observational study. *Epilepsy Res* 2014; **108**: 1817– 1824.
21. Vorderwulbecke BJ, Kowski AB, Kirschbaum A, *et al*. Long-term outcome in adolescent-onset generalized genetic epilepsies. *Epilepsia* 2017; **58**: 1244– 1250.
22. Canevini MP, Mai R, Di Marco C, *et al*. Juvenile myoclonic epilepsy of Janz: clinical observations in 60 patients. *Seizure* 1992; **1**: 291– 298.
23. Gesche J, Khanevski M, Solberg C, Beier C. Resistance to valproic acid as predictor of treatment resistance in genetic generalized epilepsies. *Epilepsia* 2017; **58**: e64–e69.
24. Kwan P, Brodie MJ. Early identification of refractory epilepsy. *N Engl J Med* 2000; **342**: 314– 319.
25. Picot M-C, Baldy-Moulinier M, Daurès J-P, Dujols P, Crespel A. The prevalence of epilepsy and pharmacoresistant epilepsy in adults: a population-based study in a Western European country. *Epilepsia* 2008; **49**: 1230– 1238.
26. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs. *JAMA Neurol* 2018; **75**: 279.
27. Szaflarski JP, Rackley AY, Lindsell CJ, Szaflarski M, Yates SL. Seizure control in patients with epilepsy: the physician vs. medication factors. *BMC Health Serv Res* 2008; **8**: 264.
28. De Haan GJ, Halley DJJ, Doelman JC, *et al*. Univerricht-Lundborg disease: underdiagnosed in the Netherlands. *Epilepsia* 2004; **45**: 1061– 1063.
29. Téllez-Zenteno JF, Hernández-Ronquillo L, Buckley S, Zahagun R, Rizvi S. A validation of the new definition of drug-resistant epilepsy by the International League Against Epilepsy. *Epilepsia* 2014; **55**: 829– 834.
30. Lamberink HJ, Otte WM, Geerts AT, *et al*. Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. *Lancet Neurol* 2017; **16**: 523– 531.
31. Lamberink HJ, Otte WM, Geleijns K, Braun KPJ. Antiepileptic drug withdrawal in medically and surgically treated patients: a meta-analysis of seizure recurrence and systematic review of its predictors. *Epileptic Disord* 2015; **17**: 211– 228.
32. Baykan B, Altindag EA, Bebek N, *et al*. Myoclonic seizures subside in the fourth decade in juvenile myoclonic epilepsy. *Neurology* 2008; **70**: 2123– 2129.
33. Syvertsen MR, Thuve S, Stordrange BS, *et al*. Clinical heterogeneity of juvenile myoclonic epilepsy: follow-up after an interval of more than 20 years. *Seizure* 2014; **23**: 344– 348.

34. Panayiotopoulos CP, Obeid T, Tahan AR. Juvenile myoclonic epilepsy: a 5-year prospective study. *Epilepsia* 1994; **35**: 285– 296.
35. Gelisse P, Genton P, Thomas P, Rey M, Samuelian J, Dravet C. Clinical factors of drug resistance in juvenile myoclonic epilepsy. *J Neurol Neurosurg Psychiatry* 2001; **70**: 240– 243.



## CHAPTER 11

# INDIVIDUALISED PREDICTION OF DRUG RESISTANCE AND SEIZURE RECURRENCE AFTER MEDICATION WITHDRAWAL IN PEOPLE WITH JUVENILE MYOCLONIC EPILEPSY: A SYSTEMATIC REVIEW AND INDIVIDUAL PARTICIPANT DATA META-ANALYSIS

Remi Stevelink, Dania Al-Toma, Floor E Jansen, Herm J Lamberink, Ali A Asadi-Pooya, Mohsen Farazdaghi, Gonçalo Cação, Sita Jayalakshmi, Anuja Patil, Prof Çiğdem Özkara, Şenay Aydın, Joanna Gesche, Christoph P Beier, Linda J Stephen, Martin J Brodie, Gopeekrishnan Unnithan, Ashalatha Radhakrishnan, Julia Höfler, Eugen Trinka, Roland Krause, EpiPGX Consortium, Emanuele Cerulli Irelli, Carlo Di Bonaventura, Jerzy P Szaflarski, Laura E Hernández-Vanegas, Monica L Moya-Alfaro, Yingying Zhang, Prof Dong Zhou, Nicola Pietrafusa, Nicola Specchio, Giorgi Japaridze, Prof Sándor Beniczky, Mubeen Janmohamed, Prof Patrick Kwan, Marte Syvertsen, Kaja K Selmer, Bernd J Vorderwülbecke, Martin Holtkamp, Lakshminarayanapuram G Viswanathan, Sanjib Sinha, Betül Baykan, Ebru Altindag, Felix von Podewils, Juliane Schulz, Udaya Seneviratne, Alejandro Viloria-Alebesque, Ioannis Karakis, Wendyl J D'Souza, Josemir W Sander, Bobby PC Koeleman, Willem M Otte, Kees PJ Braun

*Manuscript under review*

# Summary

**Background**

A third of people with juvenile myoclonic epilepsy (JME) are drug-resistant. Three-quarters have a seizure relapse when attempting to withdraw anti-seizure medication (ASM) after achieving seizure-freedom. It is currently impossible to predict who is likely to become drug-resistant and safely withdraw treatment. We aimed to identify predictors of drug resistance and seizure recurrence to allow for individualised prediction of treatment outcomes in people with JME.

**Methods**

We performed an individual participant data (IPD) meta-analysis based on a systematic search in EMBASE and PubMed – last updated on 11 March 2021 – including prospective and retrospective observational studies reporting on treatment outcomes of people diagnosed with JME and available seizure outcome data after a minimum one-year follow-up. We invited authors to share standardised IPD to identify predictors of drug resistance using multivariable logistic regression. We excluded pseudo-resistant subjects. A subset who attempted to withdraw ASM was included in a multivariable proportional hazards analysis on seizure recurrence after ASM withdrawal. The study was registered at the Open Science Framework (OSF; https://osf.io/b9zjc/).

**Findings**

Our search yielded 1641 articles; 53 were eligible, of which the authors of 24 studies agreed to collaborate by sharing IPD. Using data from 2518 people with JME, we found nine independent predictors of drug resistance: three seizure types, psychiatric comorbidities, catamenial epilepsy, epileptiform focality, ethnicity, history of CAE, family history of epilepsy, status epilepticus, and febrile seizures. Internal-external cross-validation of our multivariable model showed an area under the receiver operating characteristic curve of 0·70 (95%CI 0·68-0·72). Recurrence of seizures after ASM withdrawal (n=368) was predicted by an earlier age at the start of withdrawal, shorter seizure-free interval and more currently used ASMs, resulting in an average internal-external cross-validation concordance-statistic of 0·70 (95%CI 0·68-0·73).

**Interpretation**
We were able to predict and validate clinically relevant personalised treatment outcomes for people with JME. Individualised predictions are accessible as nomograms and web-based tools.

**Funding**
MING fonds

# Research in context

## Evidence before this study

Juvenile myoclonic epilepsy (JME) is a common generalised epilepsy syndrome. According to the 1989 definition of JME, 'response to appropriate anti-seizure medication (ASM) is good', and this is a recurring assumption. In a recent systematic review and meta-analysis, we reported that a third of people with JME are drug-resistant. We also found that three-quarters of individuals attempting to withdraw treatment after becoming seizure-free experienced a recurrence of seizures, far more than most other epilepsy types. We last updated our systematic literature search in PubMed and Embase on 11 March 2021 to identify observational cohorts describing treatment outcomes of people with JME. Previous studies identified potential predictors of drug-resistant JME but used univariable or underpowered multivariable analyses. Thus, it is unknown which variables have independent predictive power for drug resistance. There are no known predictors for seizure recurrence after ASM withdrawal in JME. Several predictors of seizure recurrence have previously been identified in broader epilepsy populations, but it is unclear if these can be generalised to JME.

## Added value of this study

This individual participant data (IPD) meta-analysis used 41 variables from 2518 people to find predictors of drug resistance and seizure recurrence in JME. We identified nine independent predictors of drug-resistant JME, seven previously reported, and two novel: ethnicity and family history of epilepsy. We found three predictors of seizure recurrence after treatment withdrawal. Only one of the previously identified risk factors of post-

withdrawal relapse in broader epilepsy populations was also predictive in JME. The strongest predictor for post-withdrawal seizure recurrence in JME – earlier age at the start of withdrawal – had an inverse direction of effect compared to other epilepsy types. We created prediction models, visualised with nomograms and a web-based calculation tool, which showed good predictive performance at the individual participant level. We performed internal-external cross-validation of the drug resistance and post-withdrawal seizure recurrence prediction models, which showed robust external predictive performance.

## Implications of all the available evidence

We created and validated prediction models, available as nomograms and web-based tools, to improve and personalise JME treatment. For example, early referral to an epilepsy clinic should be considered for people at risk of drug resistance. The risk-benefit ratio of valproate should be carefully assessed at childbearing age due to its superiority in seizure control and teratogenicity risk. Individualised prediction of a low seizure relapse risk could guide specific individuals towards a relatively safe ASM withdrawal attempt. In contrast, others – with a high-risk profile – should remain on therapy.

# Introduction

Juvenile myoclonic epilepsy (JME) is the most common idiopathic and presumed genetic generalised epilepsy syndrome, affecting 5-10% of all people with epilepsy.[1] Response to anti-seizure medication (ASM) is often assumed to be good,[2,3] but we recently reported that a third of all people with JME are drug-resistant.[4] People with JME are widely believed to require lifelong treatment.[5,6] After a period of seizure freedom, however, around a quarter of those who withdraw treatment may remain seizure-free.[4]

Predicting which individual is likely to become drug-resistant those who could safely withdraw ASM treatment after a certain period of sustained seizure-freedom, has clinical benefits. Drug withdrawal improves the quality of life by avoiding the adverse effects of potentially unnecessary treatment.[7,8] We previously identified 43 reports providing treatment outcomes in cohorts of people with JME.[4] We found six prognostic risk factors for drug resistance. Some of these risk factors are collinear, and it is unknown which have independent predictive value. Recent published multivariable prediction models of drug resistance had intrinsic limitations due to relatively small and heterogeneous cohorts, including different types of generalised epilepsy.[9,10] There are currently no known risk factors for seizure relapse after ASM withdrawal in JME, other than those previously identified in the broader epilepsy population.[11]

We aimed to identify independent predictors of drug resistance and post-withdrawal relapse risk based on individual participant data (IPD) from previously published study cohorts. We developed and validated predictive tools to calculate these risks.



# Methods

## Search strategy and selection criteria

We performed a meta-analysis of individual participant data according to a pre-registered protocol (https://osf.io/b9zjc/). The methods and reporting are consistent with the PRISMA-IPD[12] and TRIPOD statements.[13] We systematically searched PubMed and EMBASE for articles published in English, Dutch, German, Spanish or French describing treatment outcomes in observational cohorts of people with JME, with no date restrictions. The literature search

was last updated by RS on 11 March 2021, using the same search terms and study-level inclusion and exclusion criteria as in a previously published meta-analysis:[4] we included retrospective and prospective studies reporting on treatment outcomes of people with a diagnosis of JME. As individuals were diagnosed before the proposed consensus criteria,[5] our diagnoses were primarily made according to descriptive criteria. JME is a distinctive syndrome characterised by juvenile-onset myoclonic seizures and generalised tonic-clonic seizures (GTCS), usually occurring after wakening and evoked by sleep deprivation, alcohol consumption, and especially a combination of irregular spike and wave discharges in the EEG.[2] We excluded articles exclusively reporting on people with drug-resistant JME or in remission. We excluded drug trials as these could be biased towards people with drug-resistant JME. We used three individual-level inclusion criteria: (1) clinical diagnosis of JME, regardless of the diagnostic criteria used by the study,[5] (2) at least one year of follow-up, and (3) available information regarding seizure outcome, with ASM use. People with pseudo-resistant epilepsy were excluded, i.e. seizures due to non-compliance, inadequate treatment or inadequate lifestyle regulation.[14] People who had attempted to withdraw ASM after a period of seizure freedom were included in an analysis to assess predictors of seizure relapse after ASM withdrawal. We found one additional article by cross-referencing.

We invited the corresponding authors of all potentially eligible studies to collaborate by sharing IPD. If we received no reply, we sent two reminders 4-6 weeks apart, and when possible, we contacted additional authors of the same study. We searched ResearchGate, the International League Against Epilepsy (ILAE) website, other publications by the same authors and performed a manual internet search for alternative contact details.

Authors who agreed to collaborate were asked to provide treatment outcome data and potential predictors by filling in a standardised data entry sheet containing 41 variables (supplementary table 1). Alternatively, collaborators could send a datasheet in their format, after which the coordinating investigator standardised the data. Some collaborators could update their data with additional variables or subjects not included in their original publication. All datasets were manually reviewed, and potential discrepancies were resolved by discussion with the contributing author. We did not include aggregate study data without IPD.

As in our previous meta-analysis, we used the Newcastle-Ottawa quality assessment scale for cohort studies to assess the methodological quality of all the included studies.[15] The scale ranges between 1-8, where higher scores represent a higher quality and less risk of bias.

Our study was a meta-analysis of de-identified individual data and did not require ethical approval or specific informed consent. Local research ethics committees or other entities overseeing personal data had approved the original studies. Where applicable under local regulations, data sharing agreements were signed before receiving individual data.

## Outcome and predictor variables

We used a combination of outcome measures to define drug resistance and seizure recurrence after ASM withdrawal. For the primary analysis, we used the definition of drug-resistant epilepsy formulated by the ILAE, taking each seizure type into account.[16] This was defined as the failure of two or more adequate trials of well-tolerated and appropriately chosen drug schedules. We assessed whether people had not had seizures of any type in the last one, two or five years of follow-up as sensitivity analyses. Similarly, we specifically ascertained whether individuals were free of GTCS in the last one, two or five years of follow-up, as these are the most debilitating seizure type and are less likely to be underreported.

We assessed seizure recurrence in a subset of people who attempted to withdraw treatment after a period of seizure freedom. Seizure recurrence was evaluated at two and five years after initiation of ASM withdrawal. Our primary analysis comprised recurrence of any seizure after start of ASM withdrawal. We also specifically assessed GTCS recurrence.

We selected candidate predictors of drug resistance and seizure recurrence based on our previous meta-analysis on refractory JME,[4] our previous publication on seizure recurrence in general cohorts of people with epilepsy[11], and potential predictors identified by included studies. We focused on readily available predictors in a routine clinical setting, excluding variables such as advanced EEG processing data or functional MRI biomarkers. Supplementary table 1 provides an overview of all outcome measures, predictors, and definitions.

## Data analysis

The supplementary methods provide a detailed overview of all analyses and statistical methods. In brief, the proportion of drug-resistant JME was assessed with random-effects meta-analysis and a meta-regression of drug resistance by publication year. Multiple imputations were used to deal with missing data.[17] Mixed-effects logistic regression analyses were performed to evaluate potential risk factors for drug resistance. First, all predictors with $p<0{\cdot}2$ were taken forward to a multivariable model. The model was reduced by backward selection of the least contributing variables, based on the minimisation of the Akaike information criterion. Internal-external cross-validation was performed by leaving one cohort out of the training dataset and validating the model on each holdout cohort. The area under the receiver operating curve (AUC) was computed by merging the predictions of each cross-validation iteration.[18] As sensitivity analyses, we assessed the ability to predict freedom of any seizure and freedom of GTCS in the last one, two and five years of follow-up, based on the same predictors.

Cox proportional hazards analyses were performed to assess the time to seizure recurrence after start of ASM withdrawal. Univariable predictors at $p<0{\cdot}2$ were used for multivariable analyses, after which backward selection was performed to remove the least contributing predictors. We performed an internal-external validation by splitting the 18 cohorts with data on post-withdrawal seizure recurrence into three datasets of 6 cohorts, balanced on sample size. We trained the prediction model on 12/18 cohorts and assessed the external predictive value of this model on the left-out 6 cohorts, quantified with the concordance-statistic (C-statistic).[19] Such non-random internal-external validations qualify as external validations of the model.[20,21]

AUC and C-statistic values range between 0 and 1, where a value of 0.5 represents no better prediction than chance, and 1 represents perfect predictive performance. A value $<0.7$ is generally considered poor, $\geq 0.7$ is deemed acceptable, and $\geq 0.8$ is considered excellent.[22]

All statistical analyses were performed in RStudio Version 1.3.1093, using the packages: MICE, metafor, glmer, rms, coxme, rsample, purrr, survminer, tidyverse, ggplot, and survAUC.

## Nomogram and web-based risk assessment tool

To aid use in clinical practice, we converted our multivariable models to nomograms and web-based tools. The nomograms are visual representations of the mixed-effects logistic regression analysis on drug resistance and the Cox proportional hazards model on seizure recurrence within 2 and 5 years. They come with instructions to manually estimate clinical outcomes for an individual. Similarly, we translated the models into web-based tools where a user can fill in predictors to obtain the associated probability of a clinical outcome for an individual.

## Role of funder

The study's funder had no role in the study design, data collection, data analysis, data interpretation, drafting of the report or the decision to submit.

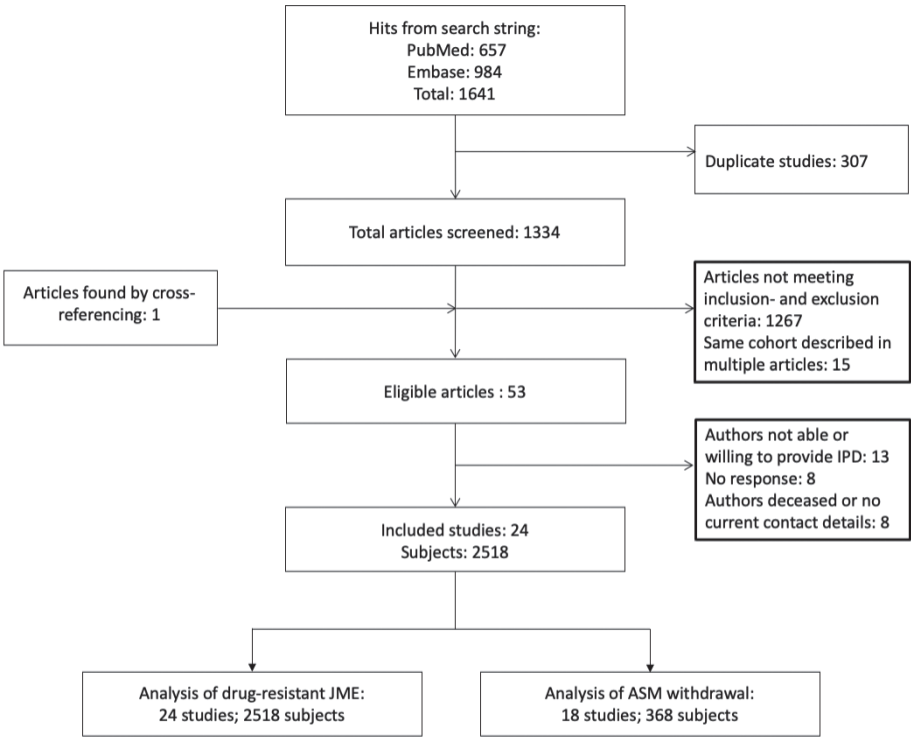



**Figure 1: Flowchart of search strategy and study selection.**

# Results

We screened 1334 articles and identified 53 eligible studies (figure 1). The authors of 24 of these studies were able and willing to provide IPD. Four were prospective, 19 were retrospective, and one study had a mixed retrospective and prospective design (see supplementary table 2 for study characteristics).[23–43] Eligible articles of which IPD was not included were similar in design and proportion of drug resistance, although some were markedly older and smaller (supplementary table 3). Meta-regression incorporating non-included articles, did not show changes in drug resistance by publication year (p=0.44; supplementary figure 1). Based on the Newcastle Ottowa assessment scale, the original publications' quality scores ranged between three to seven (mean 4·4; supplementary table 4).

In total, 2518 individuals from 18 countries and various ethnicities were included in the predictive analyses of drug resistance. Missing data before imputation ranged between 0% and 38% per variable (median 10·0%, IQR 0–19·4%). Among variables included in the drug resistance analysis, three variables had missing data between 25% and 40%; eight variables were missing between 10% and 25%, five variables between 1% and 10%, and data was complete for seven variables (supplementary figure 2). Among variables included in the seizure recurrence analysis, two variables were missing between 25 and 45%, nine variables between 10% and 25%, ten variables between 1% and 10%, and data were complete for ten variables (supplementary figure 3).

Follow-up duration ranged from one to 73 years (median 8·0, IQR 4·0–16·0). Information on current ASM treatment was available for 2365 people, of which 805 (34%) were on multiple ASMs, and 1560 (66%) were on monotherapy (supplementary table 5). Among those on monotherapy, valproate was most often used (n=826, 54%), followed by levetiracetam (n=352, 23%) and lamotrigine (n=154, 10%). Amongst 2216 people with known past ASM treatment, 661 (30%) were still taking the first prescribed medication, 727 (33%) had used one, and 828 (37%) used two or more previous ASMs (supplementary table 6). A subset of 368 people with JME (15% of the total cohort) had attempted to withdraw from ASM treatment at any time during follow-up (median follow-up after the start of withdrawal: 4·0 years, IQR 1·5–9·0). Of these, 112 (30%) were not using any ASM at the last follow-up.

Meta-analysis showed that 29% (95%CI 23-36%) of people with JME were drug-resistant (supplementary figure 4), with significant heterogeneity between studies ($I^2$=88%, p<0·0001). Amongst 388 drug-resistant people with data on the most extended period of seizure freedom, 250 (64%) had never been seizure-free for more than 12 months, and 58 (15%) were never free of seizures for more than one month at any point.

Univariable mixed-effects logistic regression analysis identified 18 predictors of drug resistance at p<0·2 (table 1; distributions of drug resistance and seizure recurrence concerning potential predictors are in supplementary table 7), some of which were correlated (see supplementary table 8). After backward selection in multivariable analyses, we identified nine variables with independent predictive values for drug-resistant JME (figure 2, supplemental Table 9): psychiatric comorbidities, three seizure types, focal epileptiform activity on EEG, catamenial epilepsy, status epilepticus, history of febrile seizures, family history of epilepsy, history of CAE progressing to JME, and ethnicity. Associations were similar when restricted to 1163 cases with complete data (supplementary table 10).

We performed internal-external cross-validation to assess the external predictive value of the multivariable model, which showed an AUC of 0·70 (95%CI 0·67-0·72). The AUC varied between 0·56 and 0·84 per left-out cohort (median 0·70, IQR 0·66-0·76), with smaller cohorts on both ends of the distribution (supplementary table 11). A plot of predicted against observed probabilities showed excellent calibration (figure 2B).

As further sensitivity analyses, we assessed how well we could predict freedom of any seizure and freedom of GTCS in the last one, two and five years of follow-up. We used the same predictors (see supplementary table 12 for a correlation matrix of outcome measures), without considering the pre-treatment seizure interval and the number and appropriateness of each drug trial as in the drug resistance analyses. The AUC for freedom of any seizure was 0·67 (95%CI 0·65-0·69) for the last year, 0·63 (0·61-0·66) for two years, and 0·59 (0·56-0·61) for five years. The AUC of the prediction model for freedom of GTCS, was 0·64 (0·61-0·67) for the last year, 0·62 (0·59-0·64) for two years and 0·63 (0·60-0·66) for five years.

| | | Association with drug resistance | | Association with seizure recurrence | |
|---|---|---|---|---|---|
| **Predictor** | **n (%) or median (IQR)** | **OR (95% CI)** | **p-value** | **HR (95% CI)** | **p-value** |
| Gender | | | | | |
| Male | 978/2518 (38.8%) | 0.84 (0.69 - 1.03) | 0.088 | 0.92 (0.71 - 1.20) | 0.55 |
| Female | 1540/2518 (61.2%) | Ref | - | Ref | - |
| Age at first seizure (years) | 15 (12-17) | 0.96 (0.94 - 0.98) | 0.00013 | 1.00 (0.97 - 1.02) | 0.79 |
| Age at last moment of follow-up (years) | 29 (23-38) | 1.01 (0.76 - 1.36) | 0.922 | 0.38 (0.27 - 0.55) | <0.0001 |
| Diagnostic delay (months) | 1 (0-3) | 1.13 (0.97 - 1.32) | 0.12 | 0.99 (0.91 - 1.07) | 0.82 |
| Ethnicity | | | | | |
| Caucasian | 1263/2054 (61.5%) | Ref | - | Ref | - |
| Asian | 510/2054 (24.8%) | 0.46 (0.24 - 0.88) | 0.022 | 0.91 (0.68 - 1.22) | 0.53 |
| Latin-American | 85/2054 (4.2%) | 1.31 (0.52 - 3.31) | 0.56 | 0.94 (0.45 - 1.96) | 0.87 |
| Other or admixed | 196/2054 (9.5%) | 1.01 (0.32 - 3.20) | 0.98 | 0.82 (0.26 - 2.61) | 0.74 |
| History of febrile seizures | | | | | |
| Yes | 205/2331 (8.8%) | 1.57 (1.14 - 2.17) | 0.0065 | 0.69 (0.42 - 1.13) | 0.14 |
| No | 2126/2331 (91.2%) | Ref | - | Ref | - |
| Ever experienced status epilepticus | | | | | |
| Yes | 68/2124 (3.2%) | 2.29 (1.37 - 3.84) | 0.0018 | 1.22 (0.50 - 2.97) | 0.66 |
| No | 2056/2124 (96.8%) | Ref | - | Ref | - |
| Developmental delay | | | | | |
| Yes | 20/2011 (1%) | 1.85 (0.76 - 4.49) | 0.18 | 1.83 (0.58 – 5.74) | 0.30 |
| No | 1991/2011 (99.0%) | Ref | - | | |
| Neurological comorbidities | | | | | |
| Yes | 190/2518 (13.5%) | 1.19 (0.84 - 1.68) | 0.33 | 1.07 (0.76 - 1.51) | 0.68 |
| No | 2328/2518 (86.5%) | Ref | - | | - |

| | | Association with drug resistance | | Association with seizure recurrence | |
|---|---|---|---|---|---|
| Psychiatric comorbidities | | | | | |
| Yes | 416/2018 (20.6%) | 2.27 (1.78 - 2.89) | <0.0001 | 1.28 (0.95 - 1.74) | 0.11 |
| No | 1602/2018 (79.4%) | Ref | - | Ref | - |
| Family history of epilepsy | | | | | |
| Yes | 796/2318 (34.3%) | 1.22 (0.99 - 1.51) | 0.068 | 0.90 (0.68 - 1.19) | 0.45 |
| No | 1522/2318 (65.7%) | Ref | - | Ref | - |
| Myoclonic seizures | | | | | |
| Yes | 2481/2518 (98.5%) | Ref | - | 0.37 (0.02 - 5.92) | 0.49 |
| No | 37/2518 (1.5%) | 0.95 (0.41- 2.18) | 0.90 | Ref | - |
| Generalised tonic-clonic seizures (GTCS) | | | | | |
| Yes | 2322/2518 (92.2%) | 1.42 (0.90 - 2.26) | 0.14 | 1.20 (0.71 - 2.03) | 0.49 |
| No | 196/2518 (7.8%) | Ref | - | Ref | - |
| Absence seizures | | | | | |
| Yes | 788/2518 (31.3%) | 2.93 (2.36 - 3.63) | <0.0001 | 1.48 (1.14 - 1.91) | 0.0031 |
| No | 1730/2518 (68.7%) | Ref | - | Ref | - |
| Three seizure types | | | | | |
| Yes | 748/2518 (29.7%) | 3.26 (2.64 - 4.02) | <0.0001 | 1.43 (1.10 - 1.86) | 0.0089 |
| No | 1770/2518 (70.3%) | Ref | - | Ref | - |
| History of childhood absence epilepsy (CAE) progressing to JME | | | | | |
| Yes | 193/2247 (8.6%) | 2.34 (1.67 - 3.29) | <0.0001 | 0.89 (0.56 - 1.41) | 0.61 |
| No | 2054/2247 (91.4%) | Ref | - | Ref | - |
| Praxis-induced seizures | | | | | |
| Yes | 103/1609 (6.4%) | 1.76 (1.05 - 2.95) | 0.034 | 1.10 (0.61 - 1.99) | 0.75 |
| No | 1506/1609 (93.6%) | Ref | - | Ref | - |
| Epileptiform focality on EEG | | | | | |
| Yes | 325/2042 (16%) | 2.23 (1.59 - 3.13) | <0.0001 | 1.19 (0.86 - 1.66) | 0.29 |
| No | 1717/2042 (84.0%) | Ref | - | Ref | - |
| Photoparoxysmal response | | | | | |
| Yes | 485/2266 (21.4%) | 1.26 (0.99 - 1.61) | 0.059 | 1.28 (0.96 - 1.70) | 0.096 |
| No | 1781/2266 (78.6%) | Ref | - | Ref | - |

| | | Association with drug resistance | | Association with seizure recurrence | |
|---|---|---|---|---|---|
| Motor seizures during sleep | | | | | |
| Yes | 266/1563 (17%) | 1.71 (1.21 - 2.41) | 0.0032 | 0.79 (0.54 - 1.17) | 0.25 |
| No | 1297/1563 (83.0%) | Ref | - | Ref | - |
| Catamenial epilepsy* | | | | | |
| Yes | 156/915 (17.0%) | 2.13 (1.44 - 3.16) | <0.0001 | 1.27 (0.82 - 1.98) | 0.28 |
| No | 759/915 (83.0%) | Ref | - | Ref | - |
| Age at start of ASM reduction (years) | 24 (19-31.75) | - | - | 0.96 (0.95 - 0.98) | <0.0001 |
| Epilepsy duration before remission (years) | 8 (3-14) | - | - | 0.97 (0.95 - 0.98) | <0.0001 |
| Seizure-free interval before start of ASM reduction (years) | 3 (2-5) | - | - | 0.58 (0.46 - 0.74) | <0.0001 |
| Number of GTCS before remission | | | | | |
| < 10 | 238/289 (82.7%) | - | - | Ref | - |
| > 10 | 51/289 (17.3%) | | | 1.03 (0.69 - 1.53) | 0.90 |
| EEG abnormality before reduction of ASM | 67/221 (30.0%) | - | - | 0.93 (0.65 - 1.34) | 0.70 |
| Yes | 154/221 (70.0%) | | | Ref | - |
| No | | | | | |
| Number of ASMs used at start of reduction | 1 (1-1) | - | - | 1.36 (1.01 - 1.82) | 0.045 |

**Table 1: Univariate predictors of drug resistance and seizure recurrence after ASM withdrawal.** The first column notes the prevalence (%) for categorical variables or the median (IQR) for numerical variables. Missing data differs per variable, and proportions are calculated based on non-missing data. Odds ratios (OR) are computed for the association with drug resistance and hazard ratios (HR) are computed to assess associations with seizure recurrence. Positive HR or OR values for numerical variables represent increased risk associated with a higher value. The last six variables are specific for subjects that have attempted ASM withdrawal thus we did not calculate associations with drug resistance. *Catamenial epilepsy is a female-specific risk factor. Thus, we have calculated the proportion as a fraction of female subjects.

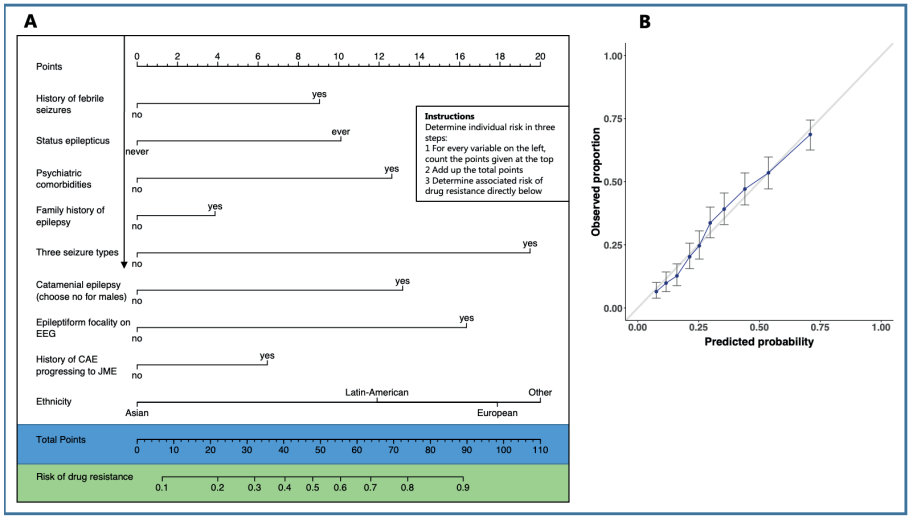



**Figure 2: (A) Nomogram for prediction of drug-resistant JME.** For example, a Caucasian (18 points) girl with a history of febrile seizures (9), who never had a status epilepticus, who has a psychiatric comorbidity (12.5), no family history of epilepsy (0), three seizure types (19.5), catamenial epilepsy (13), focal epileptiform activity on EEG (17), no history of childhood absence epilepsy (CAE; 0), has a total 89 points, corresponding to a 90% risk of drug resistance. **(B)** Calibration plot comparing observed and predicted probabilities, which should ideally follow the diagonal line.

We performed survival analyses to assess the recurrence of seizures in people who attempted to withdraw their ASM treatment (n=368). These subjects were older and included more people of Asian ethnicity compared to people who did not try to withdraw treatment but did not differ in other predictors of drug resistance (supplementary table 13). Five years after initiation of ASM withdrawal, 73% (95%CI 67-78%) had experienced seizure relapses (figure 3). Slightly fewer (69%, 95%CI 62-74%) had a seizure recurrence when assessing only GTCS. Amongst 116 people restarting treatment after a seizure recurrence and followed at least two years after recurrence, 90 (78%) regained freedom of any seizure for at least 12 months at the last follow-up.

Univariable analyses showed ten predictors of seizure recurrence at $p<0{\cdot}2$ (table 1). Subsequent multivariable analyses and backwards selection showed three variables with independent predictive value (supplementary table 14, figure 4): age at withdrawal, the seizure-free interval before withdrawal and number of ASM used at the start of reduction. Restricting analyses to 282 complete cases did not affect these associations (supplementary table 15).

Internal-external cross-validations by creating three splits of our data (6 cohorts per split) showed similar external predictive performance for all three data splits (split 1: n=119, C-statistic=0·68; split 2: n=121, C-statistic=0·74; split 3: n=128, C-statistic=0·70), with an average C-statistic of 0·70 (95%CI 0·68-0·73). Plotting observed against predicted probabilities showed good calibration (figure 4B). As an example, only 44% (95%CI 33-53%) of people older than 30 years at withdrawal had a recurrence of seizures within two years, compared to 68% (95%CI 61-73%) of those less than 30 years old (supplementary figure 5). Assessment of recurrence of GTCS after ASM withdrawal revealed one additional risk factor: people who had more than ten GTCS before remission of seizures (supplementary table 16), but the external predictive value for recurrence of GTCS was poor (split 1: n=94, C-statistic=0·64; split 2: n=94, C-statistic=0·56; split 3: n=93, C-statistic=0·61), with an average C-statistic of 0·61 (95%CI 0·58-0·63).

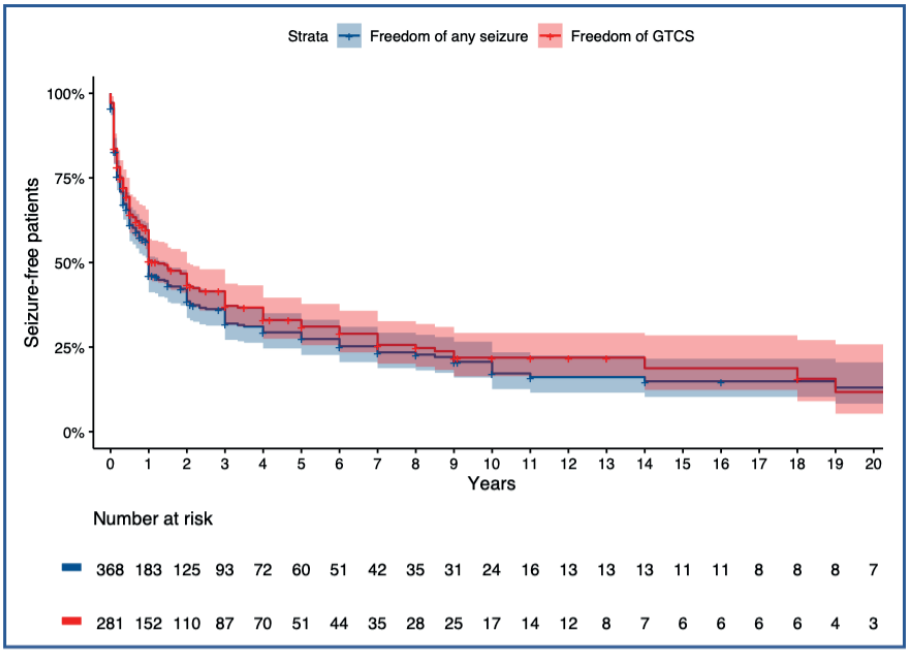



**Figure 3: Survival curve for seizure-freedom after initiation of ASM withdrawal.** Freedom of any seizure after withdrawal (blue) and freedom of generalised tonic-clonic seizures (GTCS; red) after withdrawal are displayed, with the 95% confidence interval in shaded colours. The X-axis represents the years after start of withdrawal. The number of subjects at risk is displayed below risk for each time point.

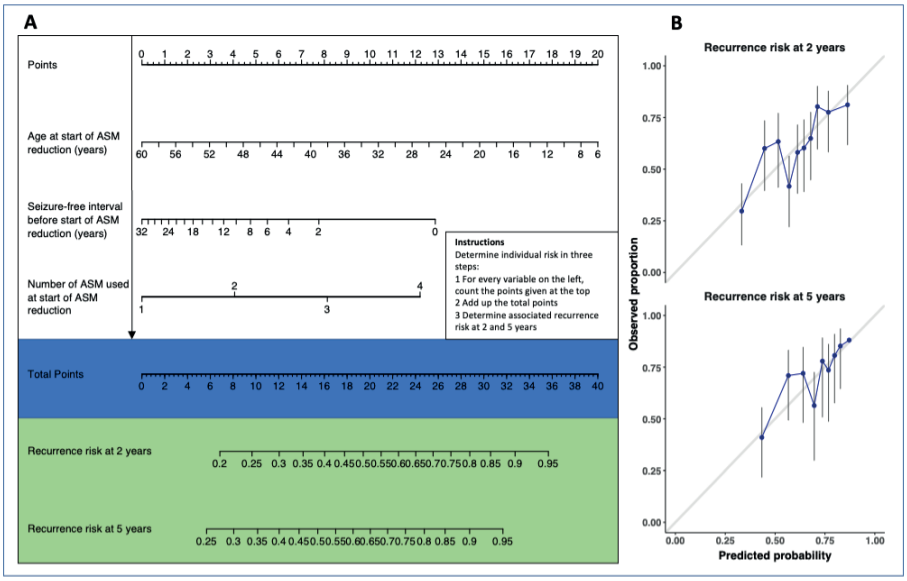



**Figure 4: (A) Nomogram for predicting recurrence of any seizure after ASM withdrawal in people with JME.** For example, someone with JME who is 44 years old at initiating withdrawal (6 points), who has been seizure-free for the last four years (6 points) and is currently using 1 ASM (0 points) has a total score of 12, which corresponds to a 37% chance of recurrence in 2 years and a 48% chance of recurrence at five years. **(B)** Calibration plot comparing observed and predicted probabilities, which should ideally follow the diagonal line.

# Discussion

We collected IPD from a large group of people, enabling the creation of prediction models for individual assessment of drug resistance (n=2518) and seizure recurrence risk after withdrawal of ASM treatment (n=368) in JME. We validated the prediction models and found good calibration. Three-quarters of people who attempted to withdraw ASMs experienced a seizure recurrence within five years, for which we found three predictors. A third of people with JME were drug-resistant, for which we found nine independent predictors.

We confirmed multiple previously found risk factors of drug resistance.[4,10,44] Some of these predictors were correlated, and we included all possible predictors in a sufficiently powered single model. We based our multivariable prediction model on routinely available variables for the model to be

freely used in clinical practice. Prediction of a relatively high risk of drug resistance could have implications for counselling and treatment guidance. For example, a Caucasian woman with catamenial epilepsy, three seizure types and psychiatric comorbidities has a high risk of drug resistance. Early referral to a specialised epilepsy clinic should then be considered. Valproate may be regarded as an option at childbearing age, but only after careful consideration of the superior efficacy versus teratogenicity.[45]

We identified two predictors not previously associated with drug-resistant JME: family history of epilepsy as a risk factor, and Asian ethnicity was protective compared to Caucasians. The most likely explanation for this association is that both predictors are proxies of the presumed population-specific genetic basis of JME.[46] Alternatively, there might be differences in under-reporting seizures relating to cultural and ethnic differences, social stigma, or driving licence regulations.[47,48]

We found three predictors of seizure recurrence after ASM withdrawal. There are several predictors of seizure recurrence in the broader epilepsy population,[11] but it is unknown whether this could be generalised to specific syndromes such as JME. We found that only one out of eight previously identified risk factors for seizure recurrence was predictive in JME. Interestingly, the strongest predictor in JME, age at ASM withdrawal, has an opposite direction of effect than in the broader epilepsy population:[11] older age at withdrawal reduces the risk of seizure recurrence in JME. In contrast, it increases seizure recurrence risk in a large population of other epilepsy forms.[11] These findings underscore the benefit of assessing a specific epilepsy syndrome instead of pooling heterogeneous epilepsy subtypes. We found that two-thirds of people with JME had a recurrence of seizures within five years, much higher than for other types of epilepsy.[11] This recurrence risk aligns with the common perception that people with JME require lifelong treatment. A subset of people, particularly older people using one ASM with prolonged seizure freedom, may have a good chance of remaining seizure-free. This is in line with the finding that myoclonic seizures often cease in the fourth decade.[25]

Our results showed a higher AUC for drug resistance prediction than seizure freedom prediction in the last one, two, and five years. The higher AUC

suggests that the formal definition of drug resistance as defined by the ILAE,[16] which considers the pre-treatment seizure interval and the number and appropriateness of each drug trial, is a more robust outcome measure than seizure freedom alone. We did not achieve excellent predictive accuracy, despite the large sample size and the checking of various independent predictors. One explanation could be that drug response may change over time, whereas the predictors remain stable throughout life.[49] Indeed, repeated remissions and relapses are common in epilepsy (although not explicitly assessed in JME-only cohorts),[50] and some people resistant to the first two ASM regimens become seizure-free upon a third or later regimen.[51,52] Conversely, two-thirds of people with drug resistance had a prior episode of remission longer than one year.[53] The ILAE definition of drug resistance outperforms previous definitions but there remains a substantial inter- and intra-observer variability.[54]

We found a higher AUC for predicting seizure freedom and recurrence of seizures after ASM withdrawal when assessing any seizure type compared to the analyses confined to GTCS. Analyses on GTCS may lack statistical power. Alternatively, freedom of GTCS might be inherently more challenging to predict since GTCS often occur less frequently than myoclonic or absence seizures. For example, assessing GTCS freedom in the last year of follow-up might be an unreliable measure for someone only having GTCS every other year. Hence, we would advocate using the ILAE definition of drug resistance, which considers all seizure types and pre-treatment seizure intervals.

Our study has limitations. The included cohorts were primarily obtained from tertiary care centres, potentially limiting the generalisability of our prediction model to primary and secondary care. Potential selection bias and selective loss to follow-up of drug-responsive people could further reduce the representativeness of our dataset. We provided a standardised data entry sheet, but significant intra- and inter-observer variability likely remain. We found considerable heterogeneity between cohorts. Potential sources of heterogeneity include differences in demography, study ascertainment, country-specific healthcare organisation and accessibility. In particular, differences in ethnicity between studies could explain part of

the heterogeneity in the proportion of drug resistance, and age differences might explain heterogeneity in seizure recurrence rate after withdrawal. We mitigated the influence of between-study heterogeneity by using random-effect statistical models, although heterogeneity might have still limited the predictive performance of our models. As most studies were several years old, collecting all potential predictors for each subject was impossible. We mitigated this by performing multiple imputations of missing data, reducing bias and increasing precision.[55] We included only readily available clinical predictors. It is possible that the predictive performance could be improved if other variables such as genetic diagnostic investigations, advanced EEG analysis, and functional MRI measures were included. Only six ethnic Asians were included in studies outside of Asia, and no Caucasians were included in Asian studies. Therefore, we were unable to perform stratified analyses on ethnicity by location. Lastly, the small proportion of subjects that attempted to withdraw treatment limited our analyses on predictors of seizure recurrence. We were unable to find an individually large enough cohort to perform external validation. However, our internal-external cross-validations performed by creating three splits of the 18 cohorts showed robust external predictive performance of our models,[20,21] suggesting that the predictors are similar across different populations. . It is essential to consider these limitations in the context of the evidence before this study. Knowledge on JME prognosis and risk factors of drug resistance is currently based on single-centre cohort studies without validation, ASM withdrawal is rarely attempted, and there are currently no known predictors to guide a safe attempt. Despite some unavoidable limitations, a meta-analysis of IPD represents the best available evidence at this moment.[56]

In conclusion, we assessed whether we could predict the likelihood that an individual with JME will become drug-resistant or has a seizure recurrence after ASM withdrawal. After validating these predictions, we created nomograms and developed publicly accessible web-based tools to help estimate individualised risks (http://epilepsypredictiontools.info/). We expect that the models will aid in improving and personalising the treatment and counselling of people with JME.

## Contributors

RS, FEJ, HJL, JWS, BPCK, WMO and KPJB contributed to study conceptualisation and design. RS and DA-T analysed the data and created the figures. WMO and HJL supervised the statistical analyses. RS and DA-T verified the integrity of the full dataset. RS, DA-T, WMO, BPCK, FEJ, KPJB have full access to all the data. RS wrote the first draft of the report, with input from DA-T, WMO, FEJ, BPCK and KPJ. Data were obtained by AAA, MF, GC, SJ, AP, ÇÖ, SA, JG, CPB, LJS, MJB, GU, AR, JH, ET, RK, ECI, CDB, JPS, LEH-V, MLM-A,YZ, DZ, NP, NS, GJ, SB, MJ, PK, MS, KKS, BJV, MH, LGV, SS, BB, EAA, FvP, JS, US, AV-A, IK, WD, JWS. All authors interpreted results, reviewed and critically revised the article, and approved the final version for submission.

## Declaration of interests

AAA received a grant from the National Institute for Medical Research Development, royalties for a book publication from Oxford University Press, and speaker fees from Cobel Daruo, Tekaje, and Raymand Rad. CPB received research grants and honoraria from UCB and Eisai, support for attending meetings by UCB, and served in the advisory board of Arvelle. CDB received consulting fees and honoraria from GW pharmaceuticals, UCB Pharma, EISAI, Angelini Pharma and Bial. JPS received grants from the National Institutes of Health, Department of Defense, and the National Science Foundation; and consulting fees from UCB Pharma, AdCel Biopharma, LLC, iFovea, SK Life Sciences, and LivaNova; and has stock options for iFovea and AdCel Biopharma. LEH-V participates in the Young Epilepsy Society, received speaker honorario from Armstrong, and was supported by Abbott pharmaceuticals to attend the Mexican Congress of Neurology. NP received honoraria from Zogenix and Ethos for Angelini Pharma. NS received honoraria from Biomarin, Livanova, GW Pharma, Zogenix and Marinus; and support to attend meetings from Livanova, GW Pharma and Zogenix; and participated on a data safety monitoring board for Marinus. SB received speaker fees from Eisai. PK received lecture honorarium from UCB Pharma and Eisai, consulting fees from Eisai and LivaNova and his institution received research grants from UCB Pharma and Eisai. MS received speaker honoraria from UCB Pharma and Eisai. KKS received research

grants from the Norwegian Research Council, the DAM Foundation and the Norwegian National Advisory Unit on Rare diseases; and a networking grant from the NordForsk Foundation; and she acted as a paid PhD defense opponent at the University of Bergen, and attended a meeting for Nordic clinicians organised by Eisai. BJV received grants from the German Society of Epileptology and the Deutsche Forschungsgemeinschaft; and honoraria from University Medical Center Schleswig-Holstein and Cornelsen Verlag. MH received consulting fees and honoraria from Arvelle, Bial, Desitin, Eisai, GW Pharmaceuticals, UCB Pharma, and Zogenix. FvP has received speaker honoraria from Bial, Eisai, GW Pharmaceutical companies, Angelinipharma, Zogenix and UCB Pharma; and scientific advisory board honoraria from GW Pharmaceutical companies, UCB Pharma, and Angelinipharma. WD's salary is part-funded by The University of Melbourne; he has received travel, investigator-initiated, scientific advisory board and speaker honoraria from UCB Pharma Australia and Global; investigator-initiated, scientific advisory board, travel and speaker honoraria from Eisai Australia and Global; advisory board honoraria from Liva Nova and Tilray; educational grants from Novartis Pharmaceuticals, Pfizer Pharmaceuticals and Sanofi-Synthelabo; educational; travel and fellowship grants from GSK Neurology Australia, and honoraria from SciGen Pharmaceuticals; and he has an equity interest in the device company EpiMinder. CPB received honaries and research support from EISAI, UCB and Arvelle. ET received speaker's honoraria from Arvelle, Abbott, Angelini Pharma, UCB, Biogen, Gerot-Lannacher, Bial, Eisai, Epilog, Takeda, Newbridge, Hikma, GW Pharmaceuticals, Sunovion Pharmaceuticals Inc., LivaNova and Novartis; consultancy funds from Angelini Pharma, Argenix, Arvelle, Epilog, UCB, Biogen, Gerot-Lannach, Bial, Eisai, Takeda, Newbridge, GW Pharmaceuticals, Sunovion Pharmaceuticals Inc., Marinus, and Novartis; directorship funds from Neuroconsult GmbH. ET's Institution received grants from Biogen, Red Bull, Merck, UCB, European Union, FWF Österreichischer Fond zur Wissenschaftsförderung, and Bundesministerium für Wissenschaft und Forschung. JWS reports personal fees from Arvelle, personal fees from UCB, grants from UCB, grants from NEF, grants from UCB, personal fees from Zogenix, grants from GW Phama, outside the submitted work; and his current position is endowed by the Epilepsy Society, he is a member of the Editorial Board of the Lancet Neurology, and

receives research support from the Marvin Weil Epilepsy Research Fund. All other authors declare no potential competing interests. None of the above mentioned declarations represent a conflict of interest directly related to the present publication.

## Data sharing

The individualised prediction models for drug resistance and recurrence of seizures after ASM withdrawal in JME are available as nomograms in this manuscript. They will be made available upon publication as a web-based tool at http://epilepsypredictiontools.info/. The main analyses' study protocol and R scripts are available on https://osf.io/b9zjc/. The signed data-sharing agreements between the different cohorts participating in this study do not allow the de-identified individual participant dataset to be publicly released. An exception can be made to replicate the results in this manuscript by an academic third party, after signing data sharing agreements with all collaborating centres.

## Acknowledgements

We are grateful to the MING fonds for supporting this project, a generous donation by parents of children with epilepsy, which provided funding for doctoral studies of RS. We would like to thank Dr. Giovanni Falcicchio (Department of Basic Medical Sciences, Neurosciences and Sense Organs, University Hospital of Bari "A. Moro", Bari, Italy) and Rachel Wales BSc (University of Glasgow) for help with data collection. The European Union Seventh Framework Programme (FP7/2007-2013) supported this work under grant agreement n° 279062, as part of the EpiPGX project. The NIH supported this work through grant NIH K23 NS052468.

## Supplementary materials

Supplementary materials can be found at: https://tinyurl.com/4j3srm5j

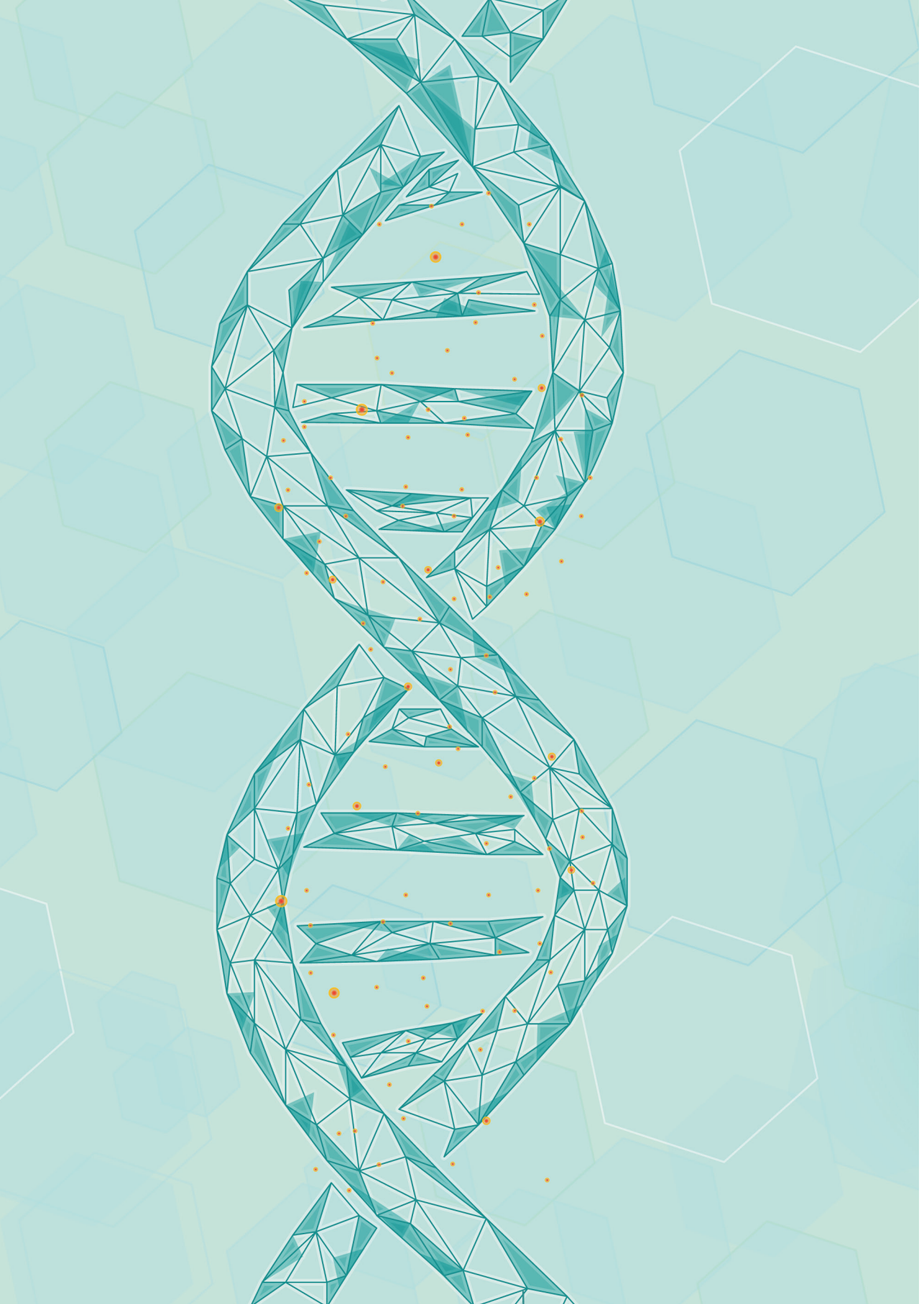
# References


1. Camfield CS, Striano P, Camfield PR. Epidemiology of juvenile myoclonic epilepsy. *Epilepsy Behav* 2013; 28 Suppl 1: S15–7.
2. Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. *Epilepsia* 1989; 30: 389–99.
3. Koutroumanidis M, Arzimanoglou A, Caraballo R, *et al.* The role of EEG in the diagnosis and classification of the epilepsy syndromes: a tool for clinical practice by the ILAE Neurophysiology Task Force (Part 1). *Epileptic Disord* 2017; 19: 233–98.
4. Stevelink R, Koeleman BPC, Sander JW, Jansen FE, Braun KPJ. Refractory juvenile myoclonic epilepsy: a meta-analysis of prevalence and risk factors. *Eur J Neurol* 2019; 26: 856–64.
5. Kasteleijn-Nolst Trenité DGA, Schmitz B, Janz D, *et al.* Consensus on diagnosis and management of JME: From founder's observations to current trends. *Epilepsy Behav* 2013; 28 Suppl 1: S87-90.
6. Vorderwülbecke BJ, Wandschneider B, Weber Y, Holtkamp M. Genetic generalized epilepsies in adults - challenging assumptions and dogmas. *Nat Rev Neurol* 2021; published online Nov 26. DOI:10.1038/s41582-021-00583-9.
7. Sillanpää M, Haataja L, Shinnar S. Perceived impact of childhood-onset epilepsy on quality of life as an adult. *Epilepsia* 2004; 45: 971–7.
8. Perucca P, Carter J, Vahle V, Gilliam FG. Adverse antiepileptic drug effects: toward a clinically and neurobiologically relevant taxonomy. *Neurology* 2009; 72: 1223–9.
9. Kamitaki BK, Janmohamed M, Kandula P, *et al.* Clinical and EEG factors associated with antiseizure medication resistance in idiopathic generalized epilepsy. *Epilepsia* 2021; published online Oct 27. DOI:10.1111/epi.17104.
10. Choi H, Detyniecki K, Bazil C, *et al.* Development and validation of a predictive model of drug-resistant genetic generalized epilepsy. *Neurology* 2020; 95: e2150–60.
11. Lamberink HJ, Otte WM, Geerts AT, *et al.* Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. *Lancet Neurol* 2017; 16: 523–31.
12. Stewart LA, Clarke M, Rovers M, *et al.* Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015; 313: 1657–65.
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015; 350: g7594.
14. Gelisse P, Genton P, Thomas P, Rey M, Samuelian JC, Dravet C. Clinical factors of drug resistance in juvenile myoclonic epilepsy. *J Neurol Neurosurg Psychiatry* 2001; 70: 240–3.

15. Wells G, Shea B, Robertson J, *et al.* The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis. 2000. http://www3.med.unipmn.it/dispense_ebm/2009-2010/Corso%20Perfezionamento%20EBM_Faggiano/NOS_oxford.pdf (accessed Dec 27, 2021).
16. Kwan P, Arzimanoglou A, Berg AT, *et al.* Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia* 2010; 51: 1069–77.
17. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011; 20: 40–9.
18. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006; 27: 861–74.
19. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011; 30: 1105–17.
20. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016; 69: 245–7.
21. Debray TPA, Moons KGM, Ahmed I, Koffijberg H, Riley RD. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat Med* 2013; 32: 3158–80.
22. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010; 5: 1315–6.
23. Viloria Alebesque A, Bellosta Diago E, Santos Lasaosa S, Mauri Llerda JA. [Juvenile myoclonic epilepsy: long-term prognosis and antiepileptic drug withdrawal]. *An Sist Sanit Navar* 2020; 43: 43–9.
24. Asadi-Pooya AA, Hashemzehi Z, Emami M. Predictors of seizure control in patients with juvenile myoclonic epilepsy (JME). *Seizure* 2014; 23: 889–91.
25. Baykan B, Altindag EA, Bebek N, *et al.* Myoclonic seizures subside in the fourth decade in juvenile myoclonic epilepsy. *Neurology* 2008; 70: 2123–9.
26. Cação G, Parra J, Mannan S, Sisodiya SM, Sander JW. Juvenile myoclonic epilepsy refractory to treatment in a tertiary referral center. *Epilepsy Behav* 2018; 82: 81–6.
27. Cerulli Irelli E, Morano A, Barone FA, *et al.* Persistent treatment resistance in genetic generalized epilepsy: A long-term outcome study in a tertiary epilepsy center. *Epilepsia* 2020; 61: 2452–60.
28. Chowdhury A, Brodie MJ. Pharmacological outcomes in juvenile myoclonic epilepsy: Support for sodium valproate. *Epilepsy Res* 2016; 119: 62–6.
29. Silvennoinen K, de Lange N, Zagaglia S, *et al.* Comparative effectiveness of antiepileptic drugs in juvenile myoclonic epilepsy. *Epilepsia Open* 2019; 4: 420–30.
30. Gesche J, Christensen J, Hjalgrim H, Rubboli G, Beier CP. Epidemiology and outcome of idiopathic generalized epilepsy in adults. *Eur J Neurol* 2020; 27: 676–84.
31. Gürer R, Aydın Ş, Özkara Ç. Outcomes of low-dose valproic acid treatment in patients with juvenile myoclonic epilepsy. *Seizure* 2019; 70: 43–8.

32. Hernández-Vanegas LE, Jara-Prado A, Ochoa A, *et al.* High-dose versus low-dose valproate for the treatment of juvenile myoclonic epilepsy: Going from low to high. *Epilepsy Behav* 2016; 61: 34–40.
33. Jayalakshmi S, Vooturi S, Bana AK, Sailaja S, Somayajula S, Mohandas S. Factors associated with lack of response to valproic acid monotherapy in juvenile myoclonic epilepsy. *Seizure* 2014; 23: 527–32.
34. Karakis I, Pathmanathan JS, Chang R, Cook EF, Cash SS, Cole AJ. Prognostic value of EEG asymmetries for development of drug-resistance in drug-naïve patients with genetic generalized epilepsies. *Clin Neurophysiol* 2014; 125: 263–9.
35. Pietrafusa N, La Neve A, de Palma L, *et al.* Juvenile myoclonic epilepsy: Long-term prognosis and risk factors. *Brain Dev* 2021; 43: 688–97.
36. Vijai J, Cherian PJ, Stlaja PN, Anand A, Radhakrishnan K. Clinical characteristics of a South Indian cohort of juvenile myoclonic epilepsy probands. *Seizure* 2003; 12: 490–6.
37. Schneider-von Podewils F, Gasse C, Geithner J, *et al.* Clinical predictors of the long-term social outcome and quality of life in juvenile myoclonic epilepsy: 20-65 years of follow-up. *Epilepsia* 2014; 55: 322–30.
38. Sun Y, Seneviratne U, Perucca P, *et al.* Generalized polyspike train: An EEG biomarker of drug-resistant idiopathic generalized epilepsy. *Neurology* 2018; 91: e1822–30.
39. Syvertsen M, Fløgstad I, Enger U, Landmark CJ, Koht J. Antiepileptic drug withdrawal in juvenile myoclonic epilepsy. *Acta Neurol Scand* 2019; 139: 192–8.
40. Viswanathan LG, Mundlamuri RC, Raghavendra K, *et al.* Long-Term Seizures Outcome in Juvenile Myoclonic Epilepsy (JME): A Retrospective Cohort Study in an Indian Population. *International Journal of Epilepsy* 2021; 7: 15–21.
41. Vorderwülbecke BJ, Kirschbaum A, Merkle H, Senf P, Holtkamp M. Discontinuing antiepileptic drugs in long-standing idiopathic generalised epilepsy. *J Neurol* 2019; 266: 2554–9.
42. Zhang Y, Chen J, Ren J, Liu W, Yang T, Zhou D. Clinical features and treatment outcomes of Juvenile myoclonic epilepsy patients. *Epilepsia Open* 2019; 4: 302–8.
43. Szaflarski JP, Lindsell CJ, Zakaria T, Banks C, Privitera MD. Seizure control in patients with idiopathic generalized epilepsies: EEG determinants of medication response. *Epilepsy Behav* 2010; 17: 525–30.
44. Chen Y, Chen J, Chen X, *et al.* Predictors of Outcome in Juvenile Myoclonic Epilepsy. *Risk Manag Healthc Policy* 2020; 13: 609–13.
45. Marson A, Burnside G, Appleton R, *et al.* The SANAD II study of the effectiveness and cost-effectiveness of valproate versus levetiracetam for newly diagnosed generalised and unclassifiable epilepsy: an open-label, non-inferiority, multicentre, phase 4, randomised controlled trial. *Lancet* 2021; 397: 1375–86.
46. 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.
47. Dalrymple J, Appleby J. Cross sectional study of reporting of epileptic seizures to general practitioners. *BMJ* 2000; 320: 94–7.

48. Amudhan S, Gururaj G, Satishchandra P. Epilepsy in India II: Impact, burden, and need for a multisectoral public health response. *Ann Indian Acad Neurol* 2015; 18: 369–81.
49. Baykan B, Martínez-Juárez IE, Altindag EA, Camfield CS, Camfield PR. Lifetime prognosis of juvenile myoclonic epilepsy. *Epilepsy Behav* 2013; 28 Suppl 1: S18–24.
50. Berg AT, Levy SR, Testa FM, D'Souza R. Remission of epilepsy after two drug failures in children: a prospective study. *Ann Neurol* 2009; 65: 510–9.
51. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment Outcomes in Patients With Newly Diagnosed Epilepsy Treated With Established and New Antiepileptic Drugs: A 30-Year Longitudinal Cohort Study. *JAMA Neurol* 2018; 75: 279–86.
52. Del Felice A, Beghi E, Boero G, *et al.* Early versus late remission in a cohort of patients with newly diagnosed epilepsy. *Epilepsia* 2010; 51: 37–42.
53. Berg AT, Vickrey BG, Testa FM, *et al.* How long does it take for epilepsy to become intractable? A prospective investigation. *Ann Neurol* 2006; 60: 73–9.
54. Téllez-Zenteno JF, Hernández-Ronquillo L, Buckley S, Zahagun R, Rizvi S. A validation of the new definition of drug-resistant epilepsy by the International League Against Epilepsy. *Epilepsia* 2014; 55: 829–34.
55. Cummings P. Missing data and multiple imputation. *JAMA Pediatr* 2013; 167: 656–61.
56. Vale CL, Rydzewska LHM, Rovers MM, *et al.* Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. *BMJ* 2015; 350: h1088.



12

**CHAPTER 12**

# GENERAL DISCUSSION AND SUMMARY

## Genetic basis of common epilepsies

Although heritability studies have long shown that epilepsy has a genetic basis, pinpointing the underlying variants and genes has long remained elusive. In our GWAS (**chapters 2 and 6**), we were able to leverage massive sample sizes through-large scale international collaborations, in order to elucidate the genetic basis of common epilepsies. We found that 39% of GGE risk can be explained by common variants (SNP-based heritability), which increases to up to 64% when assessing the subsyndrome JME, thereby largely closing the gap of missing heritability, which is estimated at 65-76%.[1–3] The remaining part of GGE liability could be due to a combination of random chance,[4] environmental factors,[5,6] and genetic variants with a larger effect, such as rare variants[7,8] and copy-number variants.[9] We observed that roughly two-thirds of the genetics variants underlying the different GGE syndromes were shared, which is reflected by the high number of genome-wide significant loci when combining the syndromes into an overall GGE GWAS. On the other hand, part of the heritability of GGE syndromes is distinct, and we identified three loci that seem specific to JME. These findings might explain why members of two-thirds of GGE families have the same syndrome, while a third of families have a mix of different GGE syndromes amongst family members.[10] We estimated that around three thousand common variants underlie GGE risk (**chapter 6**). With our current sample size, we were able to pinpoint 26 risk loci with high confidence, suggesting that we're only just seeing the tip of the iceberg. We estimated a required sample size of around 1 million to find the common variants that explain 80% of SNP-based heritability at the stringent genome-wide significance level. As a proof-of-principle, a recent GWAS of standing height including 5.4 million people was able to pinpoint 12 thousand common variants that explain almost all SNP-based heritability underlying phenotypic variation in human height.[11] Reassuringly, they found that estimates of SNP-based heritability and biological pathways from much smaller GWAS were almost identical to the much larger GWAS. Although pinpointing specific variants and genes steadily increased with sample size, these findings validate the biological insights from smaller studies like our epilepsy GWAS.

## Birth of a common variant

To understand how common variants could lead to epilepsy, it is important to understand how and why common risk variants come to exist at all. Common variants underlie most phenotypic variation between people, ranging from personality[12] and anthropomorphic traits[11,13] to risk of almost any disease.[14] Such phenotypic diversity is essential for any species to accommodate survival in a diverse and ever-changing environment. Although rare, every cell division in the body has a chance of producing a DNA mutation. Considering the quadrillions of cell divisions that occur over a lifetime, any individual is likely to accumulate some mutations. Although generally assumed to occur at random, it was only recently found that deleterious mutations occur less frequently then potentially beneficial mutations, probably due to epigenomic features, which might be an essential driving force of evolution.[15] Most variants thus arising will fade away. Only if variants are neutral, give a survival advantage, or if they piggyback an advantageous variant, can positive selection and genetic drift cause a small fraction of variants to become common over hundreds of generations.[16,17] However, if a genetic variant causes disease, how can it ever become common and why does it not fade away over time due to natural selection? To me, the most likely explanation is that disease variants also come with advantages for an individual or the population.

This is best exemplified by the archetypal psychiatric disorder schizophrenia. One might think that it is unlikely for schizophrenia risk variants to become common, since schizophrenia is associated with severely decreased life expectancy and repreduction.[18] However, family members without manifest schizophrenia, who carry part of the genetic burden, can have benefits over people without such variants. For example, they can have enhanced creativity, imagination, associative thinking and mental flexibility,[19,20] resulting in increased mating success.[21] Such variation in personality and cognitive traits might be advantageous to a wide variety in human populations, ranging from hunter-gatherer tribes to complex contemporary societies. This might explain why there are hundreds or even thousands of schizophrenia risk variants that are common in various ethnicities,[22] and why the disease remains prevalent over time.[18]

Although less understood, it is likely that common variants underlying epilepsy would similarly come with advantages. For example, unaffected siblings of people with JME – who thus have an intermediate polygenic burden – have increased brain network synchrony and functional connectivity measured on functional MRI, compared to controls without an affected family member.[23,24] Perhaps in moderation, such increases in connectivity and synchrony could have advantages for certain brain functions.

Altogether, I would advocate that the adage in toxicology "the dose makes the poison" is also applicable to common disease risk variants (Figure 1). It seems likely to me that a moderate burden of epilepsy risk variants can have advantages, whereas only an overly high burden would lead to such hypersynchrony and excitability in the brain to result in seizures. The frequency of such variants would then be determined by a balance of positive and negative selection in a population. This would explain why epilepsy and its risk variants remains common throughout the world, without fading away due to natural selection.

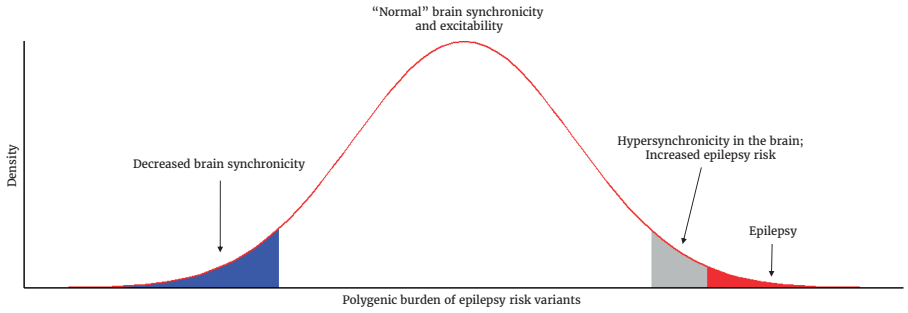



**Figure 1: A hypothesis on the relationship between polygenic burden of epilepsy risk variants, brain synchronicity and epilepsy liability.** This simplified graph illustrates my hypothesis that a low polygenic burden of epilepsy risk variants could result in decreased brain excitability, which might give a disadvantage for certain brain functions. An increased burden of risk variants could lead to hypersynchronicity in the brain, reflected by an enduring predisposition for seizures and a diagnosis of epilepsy. I expect the existence of a grey area where people with a moderately increased polygenic burden might have a moderate increased risk of epilepsy, or perhaps only experience a few sporadic provoked seizures throughout life.

## Common versus rare risk variants

It is usually thought that common variants underlie common diseases, and rare variants cause rare diseases.[25,26] Indeed, we found that the majority of GGE risk is attributable to common variants. Other studies found that rare forms of epilepsy are often caused by rare genetic variants. This can again be understood in evolutionary terms. Pathogenic rare variants are usually by themselves sufficient to cause a disease. Natural selection would therefore put strong negative pressure on such variants, preventing them from becoming common. This explains why the most severe forms of epilepsy, developmental and epileptic encephalopathies, are generally due to the rarest of variants: *de novo* mutations. Such mutations are so pathogenic that people having them are rarely healthy enough to reproduce, causing the mutation to fade away within one generation. Thus, they can only continue to exist due to spontaneous mutations in embryos or germ-line cells of unaffected parents.[27] Since such pathogenic variants are rare and bound by the size of 'susceptible' coding sequence of the disease gene and the mutation frequency, their associated phenotypes are generally also rare.

In addition to the major contribution of common variants, a large-scale exome sequencing study recently found that rare variants also contribute to various common diseases.[28] This is also true for GGE, which has a minor but non-negligible contribution from ultra-rare variants.[7,29] Copy-number variants, i.e. deletions or insertions of chunks of the genome, constitute an intermediate between common and rare variants: they are uncommon and can largely increase epilepsy risk, but they can also occur in healthy people. In a large-scale study, we found that ~3% of people with GGE carry such copy-number variants.[9]

Conversely, a surprising discovery was made in 2018 by showing that common variants also contribute to rare diseases.[30] A large cohort of children with severe neurodevelopmental disorders, assumed to be monogenic, underwent exome and common variant genotyping. They found a small but significant contribution of common variants, even in people who carried a rare pathogenic protein-coding variant. Similarly, a yet unpublished study from the Epi25 consortium showed that people with

epileptic encephalopathies have an increased polygenic burden of common variants, identified from our GWAS. It remains unclear whether such common variants influence disease severity or whether they are essential to develop epilepsy in the first place. Pathogenic epilepsy-causing variants are generally not found in healthy controls, but perhaps people with such pathogenic variants would get a different neurodevelopmental phenotype without seizures if they have a low common variant burden.

Altogether, common variants are primarily involved in common epilepsy risk and rare variants mostly underlie rare epilepsies. However, the overall emerging picture seems more complex with rare variant contributions to common epilepsies, common low risk variant contributions to rare epilepsies, and most likely everything in between. Thus, it seems likely that all genetic epilepsies are to some extent polygenic.

## Genetics of focal epilepsies

Although they are collectively common, we did not find a strong contribution of common variants for focal epilepsies. One explanation could be that focal epilepsies as a group are more heterogeneous than GGE. Some forms are more heritable than others, for example, focal epilepsy due to hippocampal sclerosis had a much higher SNP-based heritability than non-lesional focal epilepsy (**chapter 6**). Focal epilepsy could also be caused by a myriad of non-heritable causes like brain trauma, infections, stroke or tumours.[31] Furthermore, focal epilepsy by definition involves only a part of the brain. This makes it less likely that germline variants present in every brain cell would underlie focal epilepsy risk.

There are exceptions to this rule. Although present throughout the brain, some rare pathogenic variants have a predilection for certain brain areas. In fact, the first epilepsy gene to be discovered (*CHRNA4*) causes focal epilepsy.[32] Germline genetic variants in the mTOR pathway can cause focal lesions, even though they are present in each cell, which can in some cases be explained by a second-hit mechanism.[33,34] For example: a single pathogenic variant in the mTOR pathway gene *DEPDC5* may not be sufficient to cause a focal epileptogenic lesion; however, if someone has a germline

*DEPDC5* pathogenic variant in every cell, all it takes is just one cell to get a second somatic mutation in an mTOR pathway gene.[33,35] During embryonic development, there are billions of cell divisions, and in a fraction of these random mutations occur. It is likely that by chance, some of these would be in mTOR pathway genes. Since epilepsy caused by such variants involves only part of the brain, surgical resection can be very effective (**chapter 9**).

## Leveraging genetics for drug discovery and repurposing

Our increasing understanding of the genetic basis of epilepsy has led to biological insights that might aid drug discovery and repurposing. There is a great need for new drugs, considering that a third of people with epilepsy is resistant to current drugs.[36] Moreover, current drugs mostly provide symptomatic treatment of seizures and there is an urgent need for disease modifying drugs that influence or even prevent epileptogenesis.[37] Drug discovery is difficult, costly and time-consuming,[38] and most current anti-seizure medications were discovered serendipitously.[39,40] Drug repurposing based on biological insights from genetics have the potential to greatly speed up this process.[41] Drugs can be discovered by targeting single genes with high biological relevance or by assessing the aggregate effect of a large number of common genetic variations that each only explain a small part of epilepsy risk.

Treatment of tuberous sclerosis complex with vigabatrin and everolimus is a successful example of targeting a single epilepsy gene and its pathway. Tuberous sclerosis complex is caused by mutations in the mTOR pathway genes *TSC1* or *TSC2*, leading to overactivation of mTOR signalling. Everolimus is an mTOR inhibitor with proven efficacy in the treatment of epilepsy in tuberous sclerosis complex.[42] The ASM vigabatrin, which partly acts by mTOR inhibition,[43] can even delay seizures and prevent drug-resistant epilepsy when administered prior to development of clinical seizures, making it the first disease modifying drug for epilepsy in humans.[44]

In our GWAS, we found that many of the discovered epilepsy risk genes are targets of currently used ASMs, suggesting that targeting genes identified

in GWAS might be a viable approach for drug discovery. It has been estimated that selecting novel drugs based on genetic evidence can double their success rate in clinical development.[45] In **chapter 3** we investigated one such lead: the vitamin-B6 metabolism gene *PNPO*. However, our study suggested that GGE risk variants do not influence plasma levels of vitamin-B6 or its metabolites. Although we cannot rule out the possibility that *PNPO* influences brain-specific vitamin-B6 metabolism, our results do not support a role of dietary supplementation of vitamin-B6 in the treatment of GGE. Furthermore, our later GWAS (**chapter 6**), which utilized larger brain gene expression databases, suggested that a different gene in the locus (*CDK5RAP3*) might be involved instead of *PNPO*. Our larger GWAS did identify a number of novel epilepsy genes that are targeted by other drugs, which might be worthwhile to pursue for the treatment of epilepsy. For example, the novel GGE genes *CACNA2D2* and *SCN8A* are targeted by a number of currently used ASMs, but also by the Parkinson's drug safinamide, which has evidence of anti-seizure activity in animal models.[46] It should, however, be noted that since risk variants identified by GWAS individually only explain a minor contribution to epilepsy risk, it might be possible that targeting one such gene would not be sufficient to treat epilepsy. Moreover, due to the complex polygenic nature of common epilepsies, every patient would have a different combination of risk variants, making it unlikely that one drug would be effective for everyone.

Using an orthogonal approach, we assessed whether we could find candidate drugs by assessing their aggregate effect on all genes discovered by our GWAS (**chapter 5**); a method that appreciates the polygenicity of epilepsy and the polypharmacology of most drugs.[47] We intersected gene-based associations and imputed gene expressions from our GWAS with large databases that systematically assessed drug targets and their effects on gene expression. This approach produced a ranking of drugs based on their ability to target epilepsy genes and reverse abnormalities in gene expression. This ranking largely reproduced relative clinical efficacy of current ASMs, thereby validating the approach. Five highly ranked drugs not currently used for epilepsy were tested in animal models of epilepsy, of which four showed dose-dependent anti-seizure effects. Future studies could test the efficacy of these promising drug repurposing candidates in human trials.

Although drug repurposing based on genetics is an exciting and promising field, a word of caution about its potential is appropriate. Many of the current ASMs, although found serendipitously, already target the genes that were later found to be involved in epilepsy. Over the last decades many new ASMs have become available, some inspired by biological insights, but the proportion of drug-resistant people with epilepsy has not yet declined.[36] Although novel genetics-inspired drug targets could yield improvements, it seems unlikely that a miracle drug to cure all drug resistance will be found. Advances in epilepsy prediction might yield improvements within the available arsenal of epilepsy treatment and diagnosis.

## Common variants for epilepsy prediction

Although each of the common variants identified in GWAS are individually rare, they can be combined in polygenic risk scores (PRS) that can have a strong predictive value. We found that people with a high polygenic burden of GGE risk variants have a strongly increased chance of having epilepsy compared to people with a low burden (**chapter 8**). However, it remains to be elucidated whether PRS have additive predictive value above routine clinical variables. In recent years, polygenic risk scores have often been hyped[48] and even commercialised[49] to predict a wide range of diseases and traits, but substantial improvement over routine clinical predictors has not yet been demonstrated.[50,51] Ever increasing GWAS sample sizes and improved statistical methods do provide hope for PRS to improve clinical diagnosis in the future.[52,53] Moreover, the combination of PRS with clinical variables might yield a synergistic improvement in epilepsy prediction.

The hallmark of current epilepsy diagnostics is EEG measurements to detect paroxysmal epileptiform abnormalities and confirm the suggested epileptic origin of paroxysmal events. In **chapter 4** we found that common epilepsy risk variants are associated with background EEG oscillations in specific frequency bands, which are known to be highly heritable.[54] Although not yet used in clinical diagnosis, such background oscillations can easily be calculated from routinely performed EEG recordings. Further studies assessing background oscillations in people with uncertainty about the diagnosis epilepsy – for instance after a first possible seizure – could assess whether these oscillations can be used as a diagnostic biomarker.

Epilepsy is associated with widespread changes in grey and white matter, as measured on MRI, which has been considered to be a biomarker of epilepsy.[55,56] However, it is unknown whether these brain changes are cause or consequence of epilepsy. Since both epilepsy and variation in grey and white matter are strongly influenced by common genetic variation, this gives a unique possibility to leverage genetics to assess whether their causes are shared. In **chapter 7** we found that the genetic basis of common epilepsies and structural MRI measures is distinct, suggesting that the structural brain changes observed in people with epilepsy might be the consequence of epilepsy (or treatment), rather than its cause. For example, both seizure activity and the commonly used ASM valproate have been associated with grey and white matter atrophy.[57–60]

## Precision therapy through clinical prediction

Epilepsy treatment could largely benefit from personalized prediction of treatment outcomes. In **chapter 10** we describe that a third of people with JME is drug-resistant and three-quarters have a seizure relapse after medication withdrawal. We build upon this in **chapter 11** by showing that these treatment outcomes can be predicted based on a number of variables that are readily available in the clinic. Prediction of drug resistance can be used for patient counselling and to triage early referral to specialized clinics. Worldwide most people with JME are treated by primary or secondary care physicians[61] and seizure control generally improves after referral to epilepsy specialists.[62]

The question whether to continue medication usually arises in people who are seizure free while using ASM. Withdrawing medication has the potential to avoid adverse effects of unnecessary medication, but comes at the risk of seizure recurrence and limited driving license eligibility. A recent large-scale study found several predictors of seizure recurrence in a broad epilepsy population.[63] However, we found that these predictions were not generalizable to JME. Indeed, the strongest predictor – age at withdrawal – had an inverse direction of effect. These findings underline the benefits of creating prediction models based on people with the same type of epilepsy. Such prediction models have the potential to improve and personalize treatment for people with epilepsy.

There can be challenges when using prediction tools to counsel individual patients. Our prediction tools give an estimate of the risk of treatment outcomes, but it's impossible to base a decision solely on such a percentage.[64] There are various patient-specific factors that need to be taken into account. For example, when considering to withdraw ASM, it is important to note the consequences of a potential seizure recurrence on driving license eligibility, self-confidence and the individuals occupation. Moreover, it is impossible to define a universal threshold for an acceptable risk to consider ASM withdrawal. A 60% chance of seizure recurrence might seem risky to one person, whereas someone else might consider a 40% chance of remaining seizure free without any medication worth a try. Clinicians should also be careful about the framing of the assessed risk to patients, and whether to present people with a specific risk percentage. A study using an early prediction model found that presenting people with a personalized risk percentage actually made the majority of people opt to continue treatment, including people who were planning to withdraw treatment beforehand.[65] This was at least partly explained by the finding that most patients underestimated the risk of withdrawal beforehand. Framing a consultation around the risk of recurrence or emphasizing the chance of seizure freedom and its benefits could further influence the willingness to withdraw treatment.[65,66] Still, much remains uncertain about optimal counselling strategies and the real-world influence of prediction models on decision making. For such reasons, we decided not to include an arbitrary low/high risk threshold or specific recommendations based on our predictions, but we provided risk percentages that should be carefully evaluated by medical professionals in light of the individual's circumstances and preferences.



## Concluding remarks and future directions

In this thesis, we described a wide range of research projects, which I hope will aid to pave the road towards precision therapy for genetic epilepsies. Future studies might benefit from combining multiple research modalities and data sources. For example, combining rare and common variant analyses within the same cohort could establish their relative contribution.

Integration of genetics with multiple 'omics' resources can yield further biological insights and aid to prioritize drug repurposing candidates.[41] A bottleneck in many genetics studies is detailed phenotyping; often only a binary classification of disease is available for large cohorts. Combining detailed phenotyping of large cohorts with genomics could yield a bounty of results and insights. Moreover, it would facilitate assessment of the relative contribution of genetics and clinical factors to diagnose epilepsy and predict clinical outcomes. Integrating a multitude of data sources in multivariable prediction models could eventually enable precision therapy for the many people with genetic epilepsy.

Perhaps the ultimate form of precision treatment of genetic epilepsies would be the repair of causative pathogenic variants. It might be unfeasible to correct all common variants underlying polygenic GGE, but gene repair can theoretically enable curative treatment of any monogenic disease. In 2012 the CRISPR-Cas9 system, adapted from the bacterial adaptive immune system,[67] was harnessed as a toolbox for genome engineering.[68] It has since caused a revolution in the field of gene editing, recognised by a Nobel Prize in 2020. CRISPR-Cas9 is a versatile system that can be easily targeted towards any location in the genome, where it can cut and replace a stretch of DNA. Further biotechnological advances have resulted in the derived technologies prime editing[69] and base editing,[70] which can effectively repair almost any DNA alteration while causing minimal off-target mutations. This field is rapidly moving forward. A single injection of prime or base editing constructs delivered in a viral vector can already repair a sufficient amount of cells to cure monogenic diseases in mice and primates.[71,72] The first-in-human CRISPR-Cas9 clinical trial for a neurological disease showed promising efficacy and safety[73] and several more clinical trials are underway.[74] In a pilot project we were recently able to correct the causative mutation of *POLG*-related epilepsy – arguably one of the most severe and fatal forms of genetic epilepsy[75] – in patient-derived cells. Although several technological and ethical hurdles remain, I believe it is only a matter of time before gene repair becomes a feasible option for curative treatment of genetic epilepsies.

# References


1. Kjeldsen MJ, Corey LA, Christensen K, Friis ML. Epileptic seizures and syndromes in twins: the importance of genetic factors. *Epilepsy Res* 2003; **55**: 137–46.
2. Berkovic SF, Howell RA, Hay DA, Hopper JL. Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann Neurol* 1998; **43**: 435–45.
3. Corey LA, Pellock JM, Kjeldsen MJ, Nakken KO. Importance of genetic factors in the occurrence of epilepsy syndrome type: a twin study. *Epilepsy Res* 2011; **97**: 103–11.
4. Smith GD. Epidemiology, epigenetics and the "Gloomy Prospect": embracing randomness in population health research and practice. *Int J Epidemiol* 2011; **40**: 537–62.
5. Ottman R, Annegers JF, Risch N, Hauser WA, Susser M. Relations of genetic and environmental factors in the etiology of epilepsy. *Ann Neurol* 1996; **39**: 442–9.
6. Todorova MT, Mantis JG, Le M, Kim CY, Seyfried TN. Genetic and environmental interactions determine seizure susceptibility in epileptic EL mice. *Genes Brain Behav* 2006; **5**: 518–27.
7. Epi4K consortium, Epilepsy Phenome/Genome Project. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol* 2017; **16**: 135–43.
8. Motelow JE, Povysil G, Dhindsa RS, *et al.* Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. *Am J Hum Genet* 2021; **108**: 965–82.
9. Niestroj L-M, Perez-Palma E, Howrigan DP, *et al.* Epilepsy subtype-specific copy number burden observed in a genome-wide study of 17 458 subjects. *Brain* 2020; **143**: 2106–18.
10. Epi4K Consortium. Phenotypic analysis of 303 multiplex families with common epilepsies. *Brain* 2017; **140**: 2144–56.
11. Yengo L, Vedantam S, Marouli E, *et al.* A Saturated Map of Common Genetic Variants Associated with Human Height from 5.4 Million Individuals of Diverse Ancestries. bioRxiv. 2022; : 2022.01.07.475305.
12. Lo M-T, Hinds DA, Tung JY, *et al.* Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat Genet* 2017; **49**: 152–6.
13. Pulit SL, Stoneman C, Morris AP, *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* 2019; **28**: 166–74.
14. Watanabe K, Stringer S, Frei O, *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 2019; **51**: 1339–48.
15. Monroe JG, Srikant T, Carbonell-Bejerano P, *et al.* Mutation bias reflects natural selection in Arabidopsis thaliana. *Nature* 2022; published online Jan 12. DOI:10.1038/s41586-021-04269-6.
16. Charlesworth B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 2009; **10**: 195–205.

17. Lynch M, Ackerman MS, Gout J-F, *et al.* Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet* 2016; **17**: 704–14.
18. Kahn RS, Sommer IE, Murray RM, *et al.* Schizophrenia. *Nat Rev Dis Primers* 2015; **1**: 15067.
19. Mohr C, Claridge G. Schizotypy--do not worry, it is not all worrisome. *Schizophr Bull* 2015; **41 Suppl 2**: S436–43.
20. Ettinger U, Mohr C, Gooding DC, *et al.* Cognition and brain function in schizotypy: a selective review. *Schizophr Bull* 2015; **41 Suppl 2**: S417–26.
21. Nettle D, Clegg H. Schizotypy, creativity and mating success in humans. *Proc Biol Sci* 2006; **273**: 611–5.
22. The Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S, Walters JTR, O'Donovan MC. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv* 2020; : 2020.09.12.20192922.
23. Tangwiriyasakul C, Perani S, Abela E, Carmichael DW, Richardson MP. Sensorimotor network hypersynchrony as an endophenotype in families with genetic generalized epilepsy: A resting-state functional magnetic resonance imaging study. *Epilepsia* 2019; **60**: e14–9.
24. Wandschneider B, Centeno M, Vollmar C, *et al.* Motor co-activation in siblings of patients with juvenile myoclonic epilepsy: an imaging endophenotype? *Brain* 2014; **137**: 2469–79.
25. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet* 2012; **13**: 135–45.
26. Claussnitzer M, Cho JH, Collins R, *et al.* A brief history of human disease genetics. *Nature* 2020; **577**: 179–89.
27. Vadlamudi L, Dibbens LM, Lawrence KM, *et al.* Timing of de novo mutagenesis--a twin study of sodium-channel mutations. *N Engl J Med* 2010; **363**: 1335–40.
28. Wang Q, Dhindsa RS, Carss K, *et al.* Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* 2021; **597**: 527–32.
29. Epi25 Collaborative. Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. *Am J Hum Genet* 2019; **105**: 267–82.
30. Niemi MEK, Martin HC, Rice DL, *et al.* Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* 2018; **562**: 268–71.
31. Scheffer IE, Berkovic S, Capovilla G, *et al.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 2017; **58**: 512–21.
32. Steinlein OK, Mulley JC, Propping P, *et al.* A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 1995; **11**: 201–3.
33. Ribierre T, Deleuze C, Bacq A, *et al.* Second-hit mosaic mutation in mTORC1 repressor DEPDC5 causes focal cortical dysplasia-associated epilepsy. *J Clin Invest* 2018; **128**: 2452–8.

34. Baldassari S, Picard F, Verbeek NE, *et al.* The landscape of epilepsy-related GATOR1 variants. *Genet Med* 2019; **21**: 398–408.

35. Bennett MF, Hildebrand MS, Kayumi S, *et al.* Evidence for a Dual-Pathway, 2-Hit Genetic Model for Focal Cortical Dysplasia and Epilepsy. *Neurol Genet* 2022; **8**: e652.

36. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment Outcomes in Patients With Newly Diagnosed Epilepsy Treated With Established and New Antiepileptic Drugs: A 30-Year Longitudinal Cohort Study. *JAMA Neurol* 2018; **75**: 279–86.

37. French JA, Bebin M, Dichter MA, *et al.* Antiepileptogenesis and disease modification: Clinical and regulatory issues. *Epilepsia Open* 2021; **6**: 483–92.

38. Pammolli F, Magazzini L, Riccaboni M. The productivity crisis in pharmaceutical R&D. *Nat Rev Drug Discov* 2011; **10**: 428–38.

39. Nisar T, Sutherland-Foggio H, Husar W. Serendipitous antiepileptics. *Lancet Neurol* 2019; **18**: 995.

40. Nakken KO, Brodtkorb E. Chance, serendipity and antiepileptic drugs. *Tidsskr Nor Laegeforen* 2017; **137**. DOI:10.4045/tidsskr.17.0438.

41. Reay WR, Cairns MJ. Advancing the use of genome-wide association studies for drug repurposing. *Nat Rev Genet* 2021; **22**: 658–71.

42. French JA, Lawson JA, Yapici Z, *et al.* Adjunctive everolimus therapy for treatment-resistant focal-onset seizures associated with tuberous sclerosis (EXIST-3): a phase 3, randomised, double-blind, placebo-controlled study. *Lancet* 2016; **388**: 2153–63.

43. Zhang B, McDaniel SS, Rensing NR, Wong M. Vigabatrin inhibits seizures and mTOR pathway activation in a mouse model of tuberous sclerosis complex. *PLoS One* 2013; **8**: e57445.

44. Kotulska K, Kwiatkowski DJ, Curatolo P, *et al.* Prevention of Epilepsy in Infants with Tuberous Sclerosis Complex in the EPISTOP Trial. *Ann Neurol* 2021; **89**: 304–14.

45. Nelson MR, Tipney H, Painter JL, *et al.* The support of human genetic evidence for approved drug indications. *Nat Genet* 2015; **47**: 856–60.

46. Fariello RG. Safinamide. *Neurotherapeutics* 2007; **4**: 110–6.

47. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 2008; **4**: 682–90.

48. Stower H. Bringing polygenic risk scores to the clinic. Nat. Med. 2018; **24**: 1303.

49. Maxmen A. 23andMe given green light to sell DNA tests for 10 diseases. *Nature* 2017; published online April 6. DOI:10.1038/nature.2017.21802.

50. Landi I, Kaji DA, Cotter L, *et al.* Prognostic value of polygenic risk scores for adults with psychosis. *Nat Med* 2021; **27**: 1576–81.

51. Mars N, Koskela JT, Ripatti P, *et al.* Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* 2020; **26**: 549–57.

52. Wand H, Lambert SA, Tamburro C, *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *Nature* 2021; **591**: 211–9.

53. Amariuta T, Ishigaki K, Sugishita H, *et al.* Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nat Genet* 2020; **52**: 1346–54.
54. Smit DJA, Posthuma D, Boomsma DI, Geus EJC. Heritability of background EEG across the power spectrum. *Psychophysiology* 2005; **42**: 691–7.
55. Hatton SN, Huynh KH, Bonilha L, *et al.* White matter abnormalities across different epilepsy syndromes in adults: an ENIGMA-Epilepsy study. *Brain* 2020; **143**: 2454–73.
56. Whelan CD, Altmann A, Botía JA, *et al.* Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. *Brain* 2018; **141**: 391–408.
57. Galovic M, van Dooren VQH, Postma TS, *et al.* Progressive Cortical Thinning in Patients With Focal Epilepsy. *JAMA Neurol* 2019; **76**: 1230–9.
58. Coan AC, Campos BM, Yasuda CL, *et al.* Frequent seizures are associated with a network of gray matter atrophy in temporal lobe epilepsy with or without hippocampal sclerosis. *PLoS One* 2014; **9**: e85843.
59. Pardoe HR, Berg AT, Jackson GD. Sodium valproate use is associated with reduced parietal lobe thickness and brain volume. *Neurology* 2013; **80**: 1895–900.
60. Tondelli M, Vaudano AE, Sisodiya SM, Meletti S. Valproate Use Is Associated With Posterior Cortical Thinning and Ventricular Enlargement in Epilepsy Patients. *Front Neurol* 2020; **11**: 622.
61. Chen Z, Rollo B, Antonic-Baker A, *et al.* New era of personalised epilepsy management. *BMJ* 2020; **371**: m3658.
62. Szaflarski JP, Rackley AY, Lindsell CJ, Szaflarski M, Yates SL. Seizure control in patients with epilepsy: the physician vs. medication factors. *BMC Health Serv Res* 2008; **8**: 264.
63. Lamberink HJ, Otte WM, Geerts AT, *et al.* Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. *Lancet Neurol* 2017; **16**: 523–31.
64. Terman SW, Lamberink HJ, Braun KPJ. Deprescribing in Epilepsy: Do No Harm. *JAMA Neurol* 2020; **77**: 673–4.
65. Jacoby A, Baker G, Chadwick D, Johnson A. The impact of counselling with a practical statistical model on patients' decision-making about treatment for epilepsy: findings from a pilot study. *Epilepsy Res* 1993; **16**: 207–14.
66. Hux JE, Naylor CD. Communicating the benefits of chronic preventive therapy: does the format of efficacy data determine patients' acceptance of treatment? *Med Decis Making* 1995; **15**: 152–7.
67. Barrangou R, Fremaux C, Deveau H, *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007; **315**: 1709–12.
68. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012; **337**: 816–21.

69. Anzalone AV, Randolph PB, Davis JR, *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 2019; **576**: 149–57.

70. Porto EM, Komor AC, Slaymaker IM, Yeo GW. Base editing: advances and therapeutic opportunities. *Nat Rev Drug Discov* 2020; **19**: 839–59.

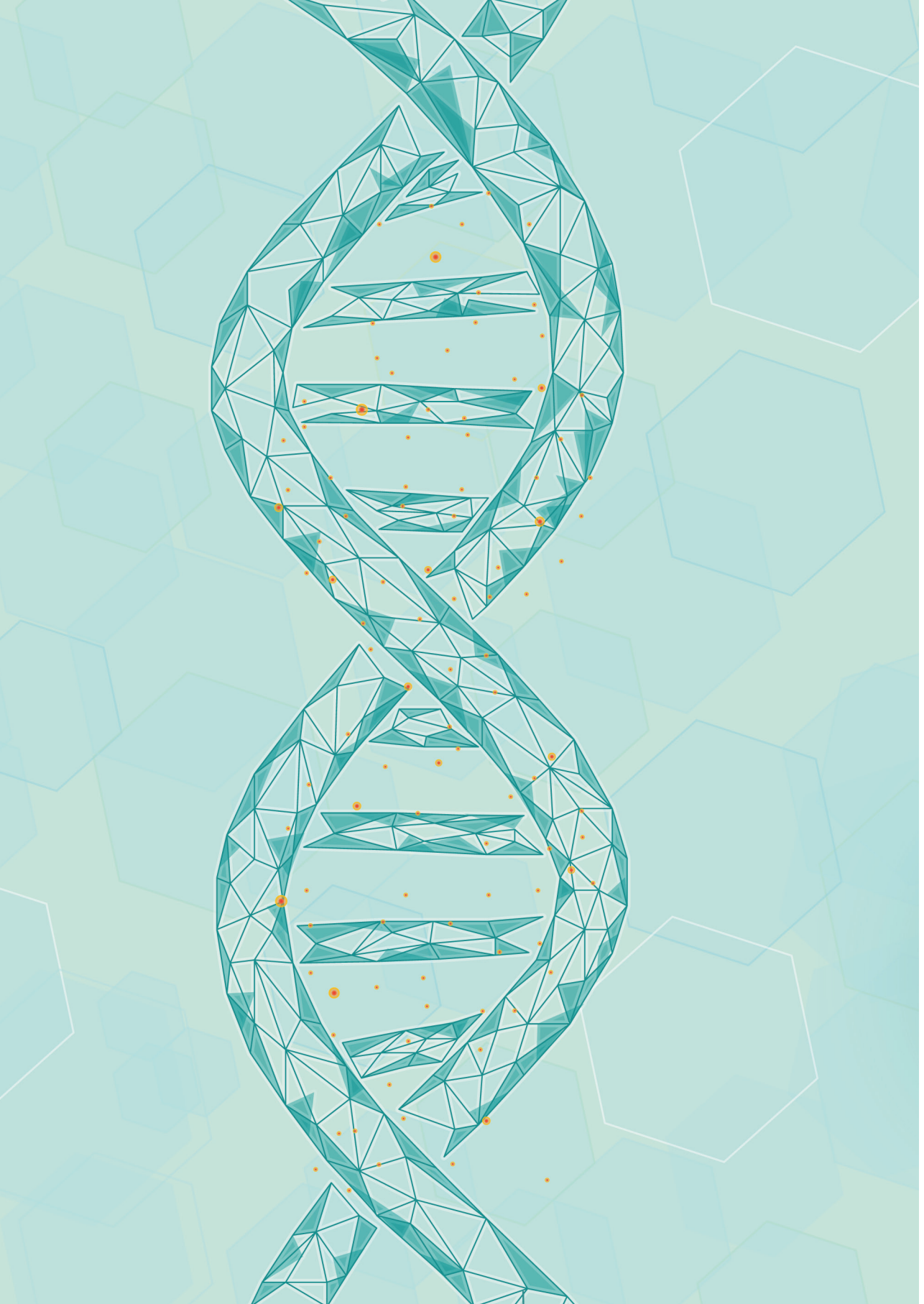
71. Musunuru K, Chadwick AC, Mizoguchi T, *et al.* In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* 2021; **593**: 429–34.

72. Jang H, Jo DH, Cho CS, *et al.* Application of prime editing to the correction of mutations and phenotypes in adult mice with liver and eye diseases. *Nat Biomed Eng* 2021; published online Aug 26. DOI:10.1038/s41551-021-00788-9.

73. Gillmore JD, Gane E, Taubel J, *et al.* CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. *N Engl J Med* 2021; **385**: 493–502.

74. Arnold C. What's new in clinical CRISPR? *Nat Med* 2021; **27**: 184–5.

75. Rahman S, Copeland WC. POLG-related disorders and their neurological manifestations. *Nat Rev Neurol* 2019; **15**: 40–52.



# APPENDIX

Nederlandse samenvatting

Abstract

Acknowledgements (Dankwoord)

List of publications

About the author

# Nederlandse samenvatting

Wereldwijd hebben zo'n 50 tot 70 miljoen de diagnose epilepsie, wat het een van de meest voorkomende neurologische ziekten maakt. Het komt voor op alle leeftijden, in elk land ter wereld en elke etniciteit. Het is al lang bekend dat epilepsie sterk overerfelijk is, waardoor we weten dat de oorzaak van veel vormen van epilepsie in het DNA is te vinden. Sinds de jaren '90 zijn er steeds meer zogenoemde epilepsiegenen ontdekt. Puntmutaties in deze genen leiden meestal tot ernstige en zeldzame vormen van epilepsie. Met DNA-onderzoek kan inmiddels voor een deel van de mensen met epilepsie een genetische diagnose worden gesteld, wat belangrijke behandelconsequenties kan hebben. Van veelvoorkomende vormen van epilepsie weten we dat deze ook overerfelijk kunnen zijn, maar vooralsnog bleek het moeilijk om hier specifieke genen voor te vinden. Met name de groep van genetisch gegeneraliseerde epilepsie (GGE), waarvan juveniele myoclonus epilepsie (JME) de meest voorkomende vorm is, komt vaak in families voor. GGE wordt niet door een mutatie in één specifiek gen veroorzaakt, maar vele genen dragen bij aan het risico op deze vorm epilepsie. Om de kleine bijdragen van individuele genetische varianten te detecteren is het noodzakelijk om het DNA van zeer grote groepen mensen met epilepsie te bestuderen. Door de genetische oorzaak van epilepsie te ontrafelen hopen we de ziekte uiteindelijk beter te kunnen begrijpen en behandelen. De huidige behandeling van epilepsie bestaat vooral uit medicijnen, maar een derde van alle patiënten blijft aanvallen houden ondanks adequaat gebruik van dergelijke anti-epileptica. Voor een deel van de mensen met focale epilepsie, epilepsie welke ontstaat vanuit een specifieke plek in de hersenen, is chirurgie een optie om epilepsie te genezen, maar voor de meeste vormen - zoals ook JME - is dit geen optie. We hopen dat de behandeling kan worden verbeterd door nieuwe behandelingen te ontwikkelen die specifieker aangrijpen op de oorzaak van de epilepsie. Binnen de huidige behandelopties is er tevens ruimte tot verbetering door beter te voorspellen wie er goed op bepaalde behandelingen reageert, en wie veilig onnodige medicatie kan afbouwen na langdurige aanvalsvrijheid.

In dit proefschrift beschrijven we de weg naar preciezere therapie voor genetische vormen van epilepsie. In **hoofdstuk 2** beschrijven we een genoom-wijde associatie studie (GWAS) waarmee we de genetische basis van veelvoorkomende vormen van epilepsie proberen te ontrafelen. Door een grote internationale samenwerking hadden we de beschikking over DNA van ruim 15 duizend mensen met epilepsie. Met deze data ontdekten we dat GGE voor

een groot deel wordt veroorzaakt door veelvoorkomende genetische varianten. Tevens konden we 16 specifieke locaties in het DNA aanwijzen die betrokken zijn bij het risico op epilepsie. Door deze GWAS-resultaten te combineren met meerdere databases en geavanceerde bioinformatische methoden toe te passen konden we uiteindelijk 21 genen aanwijzen die het meest waarschijnlijk verantwoordelijk zijn. Een hiervan was het gen PNPO, wat codeert voor het eiwit dat vitamine-B6 omzet tot zijn actieve vorm. In de studie beschreven in **hoofdstuk 3** onderzochten we of de genetische varianten verantwoordelijk voor GGE ook leiden tot een veranderde concentratie van de actieve vorm van vitamine-B6 in het bloed, wat een interessant en makkelijk therapeutisch doelwit zou kunnen zijn. Uit dit project bleek echter dat de GGE risico varianten geen invloed hebben op de hoeveelheid actief vitamine-B6 in bloed, waardoor we geen theoretische onderbouwing hebben kunnen vinden om te denken dat vitamine-B6 suppletie zou helpen voor de behandeling van GGE.

In **hoofdstuk 4** ontdekten we dat er een grote genetische overlap bestaat tussen GGE en achtergrondsignalen op het EEG. Deze overlap suggereert dat dezelfde genetische varianten leiden tot veranderde achtergrondsignalen op EEG in specifieke frequentiebanden en een verhoogd risico op GGE. Mogelijk kan dit de diagnostiek van epilepsie verbeteren, door specifieke achtergrondsignaalveranderingen in het EEG te bepalen bij mensen bij wie de diagnose niet met zekerheid gesteld kan worden.

In **hoofdstuk 5** onderzochten we of genetische bevindingen van GWAS kunnen helpen bij het ontdekken van bestaande, maar niet eerder voor epilepsie gebruikte, medicijnen om aanvallen te onderdrukken. We combineerden de epilepsie GWAS met grote databases waarin de aangrijpingspunten van medicijnen en hun effect op genexpressie verzameld zijn, om te onderzoeken welke medicijnen de processen die bij epilepsie verstoord zijn kunnen herstellen. Op deze manier konden we een ranglijst maken van middelen die het meest waarschijnlijk zouden helpen bij de behandeling van epilepsie. Geruststellend kwamen de huidige effectieve anti-epileptica hier sterk naar boven, maar we vonden ook veel nieuwe medicijnen die nog niet worden gebruikt voor epilepsie. Dit zijn middelen die al op de markt zijn voor behandeling van andere ziekten en die potentie hebben om mensen met epilepsie beter te kunnen behandelen.

In **hoofdstuk 6** beschrijven we een nieuwe GWAS met een bijna verdubbeld aantal proefpersonen: bijna 30 duizend mensen met epilepsie en 52 duizend

controles. Hiermee hadden we de mogelijkheid om nog meer risicovarianten voor epilepsie te ontdekken. Met nieuwere methoden en databases konden we 26 risicogebieden in het DNA verbinden aan specifieke genen die waarschijnlijk (mede) verantwoordelijk zijn voor het risico op epilepsie. We berekenden dat ongeveer 3000 genetische varianten samen het risico op GGE verklaren, en dat deze samen verantwoordelijk zijn voor de helft van het risico op GGE in de populatie. We hebben deze data gebruikt om nog preciezer te kunnen berekenen welke medicijnen de meeste potentie hebben om te helpen bij de behandeling van GGE.

In **hoofdstuk 7** gebruikten we de GWAS resultaten uit hoofdstuk 6 om te onderzoeken of dezelfde genetische varianten verantwoordelijk zijn voor zowel epilepsie als voor structurele veranderingen in de hersenen. Uit eerdere studies was gebleken dat de dikte van de hersenschors en de structuur van witte stof bij mensen met epilepsie anders is dan bij mensen zonder epilepsie, maar het was niet bekend of dit de oorzaak of juist het gevolg was van epilepsie. In dit hoofdstuk laten we zien dat er geen genetische overlap kan worden aangetoond, wat suggereert dat structurele veranderingen in de hersenen waarschijnlijker het gevolg zijn van herhaalde epileptische aanvallen of van de behandeling met anti-epileptica.

In **hoofdstuk 8** onderzochten we hoe bruikbaar de genetische varianten uit de GWAS in hoofdstuk 2 zijn om de diagnose epilepsie te stellen. We gebruiken zogenaamde polygene risicoscores om per persoon een score te berekenen op basis van alle varianten die geassocieerd zijn met epilepsie. Hiermee vonden we duidelijke verschillen tussen mensen met en zonder epilepsie. Dit was met name duidelijk voor GGE, maar er waren ook – hoewel minder uigesproken – verschillen tussen controles en mensen met focale epilepsie. Mensen met een hogere polygene risicoscore hadden een sterk verhoogd risico op epilepsie, maar de score lijkt nog niet sterk genoeg om voor een individu de diagnose te kunnen stellen.

In **hoofdstuk 9** beschrijven we een systematische literatuurstudie, waarin we de uitkomsten van epilepsiechirurgie voor verschillende genetische oorzaken van epilepsie bestudeerden. We vonden dat epilepsiechirurgie bij een deel van de genetische oorzaken effectief is om aanvallen volledig te bestrijden, terwijl het voor andere genetische oorzaken vrijwel nooit helpt. Deze inzichten kunnen helpen bij de selectie van kandidaten voor

epilepsiechirurgie, met name indien hun MRI scan geen afwijkingen toont. We hopen dat hierdoor enerzijds meer mensen in aanmerking kunnen komen voor een effectieve ingreep, en we anderzijds bij personen die op basis van genetische bevindingen een lage kans op chirurgisch succes hebben, invasieve diagnostiek en onnodige operaties kunnen vermijden.

Vaak wordt gedacht dat JME, net als andere vormen van GGE, goed behandelbaar is. In **hoofdstuk 10** beschrijven we echter, op basis van een systematische literatuurstudie, dat een derde van alle mensen met JME aanvallen blijft houden ondanks medicatie. Van de mensen die aanvalsvrij zijn en proberen medicatie af te bouwen kreeg driekwart weer aanvallen, hetgeen veel meer is dan voor de meeste andere vormen van epilepsie. We bouwden hier in **hoofdstuk 11** op voort door individuele patiëntendata te verzamelen van 24 cohorten, waarmee we predictiemodellen maakten voor behandeluitkomsten van JME. Met gegevens van 2518 mensen met JME konden we 9 risicofactoren vinden voor refractaire epilepsie, wat resulteerde in een statistisch model dat bruikbaar is voor individuele voorspellingen in de alledaagse praktijk. Met een tweede model konden we voorspellen wat het risico is van een terugkeer van aanvallen na het afbouwen van medicatie. We hebben deze voorspellingen beschikbaar gemaakt op een website, waarmee gemakkelijk voor individuele patiënten het risico op moeilijk behandelbare JME en het risico op terugval van aanvallen kan worden berekend. We hopen dat mensen met JME hierdoor een meer gepersonaliseerde behandeling en begeleiding krijgen.

Ten slotte plaats ik de resultaten van dit proefschrift in breder perspectief in de algemene discussie (**hoofdstuk 12**). Samengevat presenteren we een breed scala aan onderzoeken, die hopelijk kunnen helpen om preciezere therapie van epilepsie te realiseren. We komen steeds verder met het ontrafelen van de genetische basis van epilepsie. De verkregen inzichten kunnen helpen om nieuwe medicatie te ontdekken die aangrijpt op de oorzaak van epilepsie. We hopen dat deze inzichten in de toekomst leiden tot betere diagnostiek en behandeling. Onze predictiemodellen van behandeluitkomsten op basis van klinische risicofactoren kunnen nu al worden toegepast om behandelingen verder toe te spitsen op het individu. Toekomstig onderzoek kan het mogelijk maken om genetische bevindingen te combineren met klinische voorspellers, om uiteindelijk een zo persoonlijk mogelijke behandeling te kunnen bieden voor iedereen met epilepsie.

# Abstract

## English

Epilepsy is characterized by recurrent seizures and affects 50-70 million people worldwide, of which a third do not become seizure free with currently available drugs. Most current drugs are found serendipitously and despite great advances in the number of treatment options, individual treatment of people with epilepsy is currently largely trial-and-error. Many forms of epilepsy are highly heritable, in particular generalized epilepsies, suggesting that genetics could aid understanding of the pathophysiology, which might enable precision therapy aimed at the underlying cause of epilepsy. Here we describe a broad range of research techniques aimed to improve and personalize treatment of genetic epilepsies. Using large-scale genome-wide association studies, we discovered a large number of epilepsy risk variants that collectively explain much of epilepsy liability. We employed a range of methods to pinpoint the most likely implicated genes and biological processes, which we used to find drugs that target this genetic basis of epilepsy. We further assessed whether genetics could aid diagnosis or predict treatment outcomes for people with epilepsy. Finally, we created prediction models based on clinical variables that could be used to personalize treatment and counselling of people with juvenile myoclonic epilepsy, the most common form of genetic epilepsy. We hope that these findings aid to pave the road towards precision therapy of genetic epilepsies.

## Nederlands

Wereldwijd zijn er 50-70 miljoen mensen met epilepsie, waarvan een derde niet aanvalsvrij wordt met huidige medicatie. De meeste epilepsie medicijnen zijn bij toeval ontdekt en ondanks een grote toename in behandelopties blijft een groot deel van de behandeling 'trial-and-error.' Veel vormen van epilepsie, met name gegeneraliseerde vormen, zijn sterk overerfelijk, wat suggereert dat genetica kan helpen om de oorzaak te ontrafelen en om behandelingen te ontwikkelen welke op deze genetische basis aangrijpen. In dit proefschrift beschrijven we een breed scala aan onderzoek met als doel om de behandeling van epilepsie te verbeteren en personaliseren. Met genoom-wijde associatie studies hebben we veel nieuwe genetische varianten gevonden die samen het risico op epilepsie grotendeels verklaren. Met verdere analyses konden we de specifieke genen en processen welke betrokken zijn bij de genetische oorzaak van epilepsie ontrafelen, wat we hebben gebruikt om medicijnen te zoeken welke hierop aangrijpen. Vervolgens hebben we onderzocht of genetica kan helpen in het diagnosticeren van epilepsie of het voorspellen van behandeluitkomsten. Tenslotte hebben we een voorspelmodel gemaakt voor behandeluitkomsten van de meest voorkomende genetische vorm van epilepsie, juvenile myoclonus epilepsie, wat gebruikt kan worden om voorlichting en behandeling te individualiseren. We hopen dat dit onderzoek helpt om de weg vrij te maken om uiteindelijk precisie-behandeling te kunnen bieden voor mensen met genetische vormen van epilepsie.

# Acknowledgements (Dankwoord)

Dat er maar één naam op de cover van dit proefschrift mag staan doet geen recht aan de vele mensen die een cruciale bijdrage hebben geleverd. Dit proefschrift zou niet mogelijk zijn geweest zonder de hulp van mijn begeleiders, collega's, vrienden, familie en patiënten. Ik realiseer me dat er te veel mensen zijn om allemaal te benoemen, al wil ik hier toch een poging wagen.

**Dr. Koeleman, beste Bobby**, wat een genot om jou als co-promotor te hebben gehad tijdens mijn PhD. Ik kwam er al snel achter dat het een uitdaging is om specifieke meetings met je te plannen, maar gelukkig was dat ook totaal niet nodig. Ik kon altijd bij je binnenlopen of je bellen om te overleggen. We hebben een mooie synergie waarbij we urenlang kunnen discussiëren, en pas aan het einde realiseren dat we nog niet zijn begonnen met het onderwerp waar ik voor kwam. Ondanks, of wellicht dankzij, deze wat chaotische werkstijl hebben we samen enorm veel mooie en interessante projecten afgerond. Bovendien wist je me waar nodig af te remmen en tot zinnen te komen. Ik vind je enthousiasme en onuitputtelijke bron van ideeën inspirerend en bewonderenswaardig.

**Professor Braun, beste Kees**, bedankt voor al je enthousiasme en vertrouwen. Ik zie je als een groot voorbeeld en bewonder hoe je in zowel kliniek als onderzoek uitblinkt, een grote groep van collega's aanstuurt en altijd de juiste balans weet te vinden tussen daadkracht en gezelligheid. Je hebt me veel vrijheid gegeven, maar ook essentiële sturing, waardoor ik het idee heb het maximale uit mijn promotietraject te hebben gehaald. Ook heb ik er vol vertrouwen in dat ik onder jouw hoede het beste uit mezelf kan halen om me te ontwikkelen tot kinderneuroloog.

**Dr. Jansen, beste Floor**, toen ik in 2016 bij je aanklopte om te praten over een onderzoeksstage voor geneeskunde had ik nooit verwacht dat dit zou uitgroeien tot het huidige promotieonderzoek. Ik was al snel onder de indruk van je kritische en analytische blik, waarbij je vaak binnen no time met oplossingen kwam om analyses en papers te verbeteren. Bedankt voor al je steun en wijsheid gedurende de jaren. Ik weet zeker dat je het fantastisch zal doen als medisch hoofd van de kinderneurologie, en ik hoop nog lang samen te blijven werken als collega's.

Bedankt aan het **MING fonds** voor jullie enorm genereuze donatie, waardoor niet alleen het onderzoek in dit proefschrift, maar ook het onderzoek van meerdere collega's mogelijk is gemaakt. Dankzij jullie cruciale bijdrage wordt precisie-behandeling van epilepsie langzaam maar zeker realiteit.

Het onderzoek in dit proefschrift zou niet mogelijk zijn geweest zonder de onmisbare bijdrage van **duizenden mensen met en zonder epilepsie, die de tijd en moeite hebben genomen om aan wetenschappelijk onderzoek mee te doen**. Het meeste onderzoek zal niet snel genoeg verlopen om hier direct zelf baat bij te hebben, maar zonder jullie hulp zou het onmogelijk zijn om de behandeling van epilepsie verder te verbeteren.

De leden van de beoordelingscommissie, **professor Jan Veldink, professor Maeike Zijlmans, dr. Eva Brilstra, professor Edward Nieuwenhuis en professor Barbara Franke**, hartelijk dank voor uw belangstelling in mijn proefschrift en de bereidheid om zitting te nemen in de beoordelingscommissie.

The work in this thesis would not have been possible without the invaluable collaborations with a great number of colleagues. There are too many people to name everyone individually, but some colleagues I would like to thank in specific. **Professor Berkovic, dear Sam**, I have deep respect for the way you brought epilepsy genetics researchers worldwide together in the ILAE Consortium on Complex Epilepsies and the Epi25 collaborative. You have created the largest epilepsy research collaboration, which stands out in its inclusivity and openness, and which you chair with great leadership while creating an atmosphere of equality amongst all members. **Professor Cavalleri, dear JP**, thank you for your great work and guidance, which was pivotal to have the latest epilepsy GWAS come to fruition. You chaired the analysis committee with great care and eye for detail, and you made everyone feel comfortable to say whatever they thought. **Dr McCormack, dear Mark**, I was very lucky to start my PhD with you as an experienced post-doc by my side. It was great to work on the GWAS together and to be able to share and discuss every result within minutes. It was perfect to end this with great nights out in Bangkok**. Dr Lal and Dr Leu, dear Dennis and Costin**, thanks for the great collaborations on the epilepsy PRS projects, as well as good times with drinks in Bangkok. **Dr Mirza**, **dear Nasir,** thanks for all your work on translating the GWAS to meaningful drug repurposing candidates. **Professor**

**Sander, dear Ley**, I would like to thank you for sharing your wisdom and expertise, that has greatly improved our manuscripts and has helped me to navigate the complex landscape of epilepsy research. **Professor Ingrid Scheffer, Dr Karen Oliver, Dr Ciaran Campbell, Dr Siwei Chen, Dr Roland Krause, Professor Melanie Bahlo, Professor Doug Speed, and all other members of the ILAE GWAS consortium,** you have made it clear that research is teamwork. It was an honour to collaborate with such skilled, smart and warm people. I truly appreciate the welcoming and egalitarian atmosphere within the consortium, where competitiveness is absent and great things are achieved by working together as a team in harmony. **Dr Otte, beste Wim**, het statistische orakel van de epilepsie onderzoekers. Bedankt voor al je hulp met het JME predictiemodel. Zonder jouw begeleiding en expertise was dit project nooit gelukt. Ook heb je een prachtige website gemaakt, waarmee alle predictiemodellen makkelijk door internationale collega's te gebruiken zijn. **Dr Jurjen Luyxk, beste Jurjen,** bedankt voor de fijne samenwerking met het EEG project. Dankzij jouw hulp hebben we hier in snel tempo een mooie paper van gemaakt. I would like to thank **all collaborators and co-authors of the papers included in this thesis**. There are too many people who have contributed to thank everyone individually, but I think this clearly shows that research is impossible alone, and only by collaborating can we ever make progress to improve treatment and realise precision therapy for epilepsy.

Even though I eventually ended up in a different field of research, I would like to thank my previous research supervisors for teaching me the fundamentals of science, which has helped me throughout my career. **Professor Johansen-Berg, dear Heidi**, thank you for your great supervision and mentoring for my MSc research project on sleep and white matter. Although it has taken a while, I am glad to see that our persistence to get it published finally seems to be paid off. You show great leadership and the ability to connect people; the Wellcome Centre for Integrative Neuroimaing is lucky to have you as director. **Dr. Becker, dear Esther**, thank you for your guidance on my first MSc project on *CNTNAP2*. I truly appreciate your patience and guidance throughout the project, and your supervision taught me the fundamentals of basic and laboratory research. **Professor Vinkers, beste Christaan**, toen je me met mijn bachelor stage begeleidde was jij mijn eerste voorbeeld van een arts-onderzoeker. Ook toen ik tijdens mijn

geneeskunde opleiding bij je aanklopte stond de deur meteen weer open en hebben we samen twee mooie projecten afgerond. Ik ben onder de indruk hoe je geneeskunde, farmacie en rechten hebt weten te combineren, en na je opleiding in sneltreinvaart zowel psychiater als hoogleraar bent geworden.

De vele collega's met wie ik in Utrecht heb samengewerkt hebben ervoor gezorgd dat ik niet alleen veel geleerd heb, maar ook enorm genoten en veel gelachen.

**Wout**, bedankt voor de vele koffiepauzes en je hulp met het prime editen. Ook kijk ik met veel plezier terug op de Muay Thai wedstijden in Bangkok. **Joep**, ontzettend leuk dat je ons team sinds het laatste jaar hebt versterkt. Ik hoop je Zuid-Amerikaanse muziek pubquiz snel nog eens in het echt te doen. **Martijn**, de groep heeft enorm geluk om jou als onderdeel van het lab te hebben. Dankzij je harde werken en oog voor details zal de groep het ver schoppen. **Anja**, bedankt voor al je hulp met het maken van de epilepsie genetica databases. Zonder jouw hulp zou veel van het onderzoek in deze paper niet mogelijk zijn geweest. **Ruben**, de Koeleman groep heeft aan jou een stabiele basis en constante factor, en het wordt gewaardeerd dat je altijd paraat staat om collega's te helpen. **Flip**, jouw hulp om op gang te komen met mijn onderzoek en me wegwijs te maken in de (bio)informatica was onmisbaar. Bedankt voor al je hulp!

**Victoria**, super leuk om jou als achterbuur en mede onderzoeker te hebben in de groep. Ook bedankt voor het opnieuw leven inblazen van de Epilepsy Research meetings. **Lotte**, bedankt voor alle gezelligheid; jammer dat je nu naar Amsterdam bent gegaan, maar ik heb er vol vertrouwen in dat je daar een uitstekende neuroloog zal worden! **Herm**, bedankt voor al je hulp met het JME predictiemodel. Dankzij al jouw hulp met R en statistiek hoefde ik niet opnieuw het wiel uit te vinden. Je bent de bijnaam 'prediction Herm' meer dan waardig! **Bart**, na jou zal de Fledermaus in Wenen nooit meer hetzelfde zijn. Bedankt voor al je enthousiasme, en ik hoop later als kinderneuroloog nog lang collega's met je blijven. **Maurits**, ondanks dat je meestal casually late bent is zonder jou en je föhn geen enkel congres, Babinski of assistentenweekend compleet. **Hanna**, enorm leuk om jou zowel als AIOS-buddy en collega onderzoeker te hebben. Bedankt voor al je gezelligheid, hulp en wijsheid! **Carmen**, wat cool dat we tegelijk in opleiding zijn gekomen en nu samen onderzoek doen! Ik heb

er vol vertrouwen in dat team Steve nog mooie dingen gaat bereiken. **Manja**, ik waardeer enorm hoe je je inzet om de groep te verbinden. Je bent al snel een cruciale schakel in de groep worden. **Trude**, enorm leuk om jou tijdens mijn stage op Panda tegen te komen en nu samen te werken in het onderzoek aan het FSC cohort. **Crista**, super gezellig om jou niet alleen als AIOS maar nu ook als collega onderzoeker te hebben. Je enthousiasme tijdens o.a. borrels en Babinski's zal de groep zeker verder brengen.

**Rozemarijn, Edith, Glen, Iris, Helen, Karen en Sanne**, bedankt voor alle gezelligheid en koffiepauzes in het stratenum tijdens het eerste jaar van mijn onderzoek. De koffiebel blijft een van mijn beste verjaardagcadeaus. Ook de fright night in Walibi met jullie was fantastisch.

**Colleagues from the Center for Molecular Medicine (CMM)**, I would like to thank you all for the many lunch and coffee breaks and loads of fun at research retreats!

**Dr. Fuchs, beste Sabine**, ik waardeer het enorm dat je Gautam en mij de vrijheid gaf om aan het *POLG* project te werken en ik voelde me meteen thuis in je onderzoeksgroep. Waanzinnig om te zien dat je groep zo snel uitbreidt. Ik denk dat je onderzoek de behandeling van metabole ziekten radicaal zal veranderen. Ondanks dat er nog een lange weg te gaan is, heb ik er vol vertrouwen in dat we in de toekomst ook *POLG* kunnen behandelen met prime editing. **Imre en Indi**, bedankt al jullie hulp en geduld om me de basis van prime editing en labwerk te leren. Ik ben onder de indruk hoe jullie het prime editing in Utrecht hebben opgezet en zo snel internationaal toonaangevend onderzoek hebben verricht. **Eline**, zonder jouw hulp was het POLG project niks geworden. Ondanks dat je geen enkele labervaring had, werd je hier al snel beter in dan ikzelf. **Irena**, bedankt voor je hulp met de POLG organoids, en voor de borrels en feestjes.

Bedankt aan alle **collega assistenten neurologie**. Ik ben enorm blij om onderdeel te zijn van deze mooie groep. Ik vind het met iedereen enorm leuk om samen te werken, maar bovenal is het ook altijd gezellig. Dankzij jullie is er altijd een fijne sfeer en ik kijk nu al uit naar de rest van de assistentenborrels, weekenden en Babinski vakanties! Bij de volgende Babinski zal mijn pols het hopelijk meer dan 1 piste overleven.

Graag wil ik mijn opleiders vanuit de neurologie **professor Geert Jan Biessels en professor Tatjana Seute** bedanken voor alle hulp en steun tijdens mijn traject. **Beste Geert Jan en Tatjana,** jullie hebben een fijne en stimulerende leeromgeving gecreëerd, waardoor ik het idee heb het beste uit mezelf te kunnen halen. Dank ook aan **alle neurologen en kinderneurologen** met wie ik de laatste jaren heb samengewerkt in het UMCU, WKZ, Tergooi en St. Antoniusziekenhuis, ik heb van ieder van jullie veel geleerd.

Tot slot wil ik mijn vrienden en familie bedanken.

**Rein**, toen we samen dagenlang op een kameel door de Sahara in Marokko trokken en oneindig lange discussies hadden, hadden we het idee samen de wereld aan te kunnen en alles te kunnen ontdekken en begrijpen. We konden over alles goed praten, en hadden onder andere een gedeelde passie voor filosofische onderwerpen, neurologie en neurowetenschap. We hadden al snel het plan bedacht om samen neuroloog en neurowetenschapper te worden. Ik ben er trots op dat ik dit plan namens ons heb kunnen voortzetten, maar het is een enorm gemis om jou hierbij niet aan mijn zijde te hebben. Ik ben je voor altijd dankbaar voor alle mooie herinneringen.

Mijn brudahs **Lennart en Ran**, wat een geluk dat ik jullie al als kleine Remi ben tegengekomen. Al meer dan 20 jaar lachen we ons samen kapot, hebben we de beste gespreken, beleven we de mooiste avonturen en de vetste feesten. Laten we voor altijd de traditie voortzetten om elkaar voor elke verjaardag een vakantie of een vette activiteit cadeau te geven (al lopen we er al wel een paar achter). Ik weet zeker dat we voor de rest van ons leven enorm veel plezier zullen beleven en op elkaar kunnen bouwen. Ik heb enorm mazzel om via jullie nu ook vrienden te zijn met **Jas en Suus. Len**, wat was het waanzinnig om met jou op ons 18e een jaar naar alle hoeken van de wereld te maken, wat we al hadden bedacht toen we 11 jaar oud waren. Ik ben er trots op dat je inmiddels je eigen bedrijf hebt. Je ontwerp op de cover van dit proefschrift bewijst overigens dat het afmaken van een opleiding nergens voor nodig is. **Ran**, wat is het toch genieten om met jou te feesten, raven, bergen te beklimmen, op de bank uit te brakken met pannenkoeken en om heerlijk niks te doen. Ik weet zeker dat er nog eindeloze momenten zijn waarop we ons samen gaan vermaken.

**Rense en Anton**, wat fantastisch dat we na een jaar roeien met ons team 'Blauwe tandenborstel' nog steeds zulke goede vrienden zijn gebleven. We kunnen ons altijd samen kapot lachen, hebben veel mooie avonturen beleefd samen en ik heb er vol vertrouwen in dat er nog velen zullen volgen. **Rense,** waarschijnlijk ben je de meest productieve rapper en dichter, ik vind het waanzinnig dat je zelfs over een standaard avond met unit pasta nog een rap weet te schrijven. **Anton**, de beste kok van unit-pasta's, wat hebben we samen een top vakanties beleefd! Jammer dat je niet bij mijn PhD verdediging kan zijn, maar een fietsvakantie naar het Midden-Oosten lijkt me een acceptabel excuus. Het is ook een enorm plezier dat ik via jullie ook vrienden ben met **Anouk**, **Marjan, Marleen en Aldo** en ik hoop dat we nog vele jaren met de beesten gaan eten en drinken.

Huize Frits; **Bob, Carmen, Koen, Biz, Rik, Maud, Roos, Luc, Len, Timo, Ellen, Claudia, Fadia, Boris, Wouter, Peter, Joppe, en alle aanhang**. De vele jaren dat ik met jullie heb samengewoond waren fantastisch. Elke dag was het feest en gekte, maar ik kon ook altijd terecht bij jullie voor serieuze gesprekken. Jullie voelen nog steeds als familie en ik geniet ervan dat we elkaar nog vrijwel wekelijks zien. Ik heb er vol vertrouwen in dat we dit gaan doorzetten tot we later samen in bejaardenhuize Frits terechtkomen.

**Jordi**, ik had nooit verwacht dat we nog steeds vrienden zouden blijven, toen we je in 2009 op een dronken avond in Beijing tegenkwamen. Sinds de eerste keer dat we bij jou en Benthe nieuwjaar vierden in Roemenië was het al meteen duidelijk dat dit een traditie zou worden. Sindsdien zijn we met steeds meer, en ik hoop dat we deze traditie met o.a. **Ran, Len, Juul, Bob, Gautam, Simone, Max, Anton en Jas** nog vele jaren in stand houden!

**Gautam en Bob**, met jullie aan mijn zijde als paranimfen weet ik zeker dat mijn verdediging en feest een groot plezier worden. **Bob**, mijn allerbeste achter-achter-achter-achterneef, ik geniet van alle avonden met speciaalbier of whiskey, waarin we tot diep in de nacht al filosoferend proberen de wereld te verbeteren. Ik vind het waanzinnig hoe je onder het motto "20% vision, 0% fear" bergen beklimt, bergen af skiet of in een Duster rijdt; alle vakanties met jou zijn een groot feest. Ik zal mijn best doen om bij volgende hike vakanties routes te zoeken met minder afgronden. **Gautam**, ontzettend vet om met jou samen onderzoek te doen, maar ook de mooiste avonturen te beleven in

o.a. Borneo, Georgië en Kirgizië. Je hebt een aanstekelijk enthousiasme om te chillen, lachen en koffiepauzes te houden, maar vreemd genoeg zijn we ondanks dat samen ook nog eens enorm productief. Laten we de traditie in stand houden om beursaanvragen vooral te schrijven in het buitenland met speciaalbier. Op deze manier moet het toch wel een keer lukken om geld te krijgen voor ons *POLG* project!

Graag wil ook al mijn lieve neven, nichten, ooms en tantes van **de familie Stevelink en familie Rijk** bedanken. Ik voel me al mijn hele leven bij jullie thuis en bij elk feestje kan ik op jullie gezelligheid rekenen. Ook heb ik mazzel met **Roland, Trudi, Mathijs en de rest van alle Pouwen** als schoonfamilie, bij jullie is het altijd gezelligheid en lachen. Ook is het volume bij familiefeesten onovertroffen.

**Anja en Arnold, lieve pap en mam**, zonder jullie was ik er nooit geweest, laat staan dit proefschrift. Jullie hebben me de fijnste en meest liefdevolle opvoeding gegeven die ik me kan voorstellen. Dankzij jullie heb ik een zorgeloze jeugd gehad en heb ik mezelf kunnen ontwikkelen. Jullie hebben me altijd weten te stimuleren en motiveren om de juiste keuzes te maken, waardoor ik geniet van mijn leven. Bedankt voor alles! **Elwin**, wat een mazzel om jou als grote broer te hebben! In veel aspecten lijken we enorm op elkaar, inclusief motoriek en dansmoves, maar we kunnen elkaar ook goed aanvullen. Bedankt voor al het lachen, goede gesprekken en een fantastische gezamenlijke afstudeerborrel inclusief spectaculaire grachtenrace en optreden van Def Rhymz. **Lisa**, bij jou als grote zus kan ik mijn hele leven al terecht voor steun en gezelligheid. Toen we een jaar of 4 en 10 oud waren hadden we samen ons eerste boek geschreven over tante Sjaan Banaan. Ik vind het een eer om jou jaren later in dit boek te kunnen bedanken! Zonder jou waren ook mijn lievelingsnichtje **Jaël** en mijn lievelingsneefje **Isaiah** er niet geweest, met wie het altijd pret is. We gaan snel nog eens samen naar Monkey Town!

Lieve **Juul**, wat een geluk dat ik jou heb ontmoet. Alles wat ik met jou samen doe is leuk, variërend van verre reizen maken, wielrennen, feesten, lachen en speciaalbier drinken tot heerlijk niks doen op de bank met onze kat Chouffe. We lijken in veel dingen op elkaar, maar weten elkaar ook perfect aan te vullen. Zonder al jouw steun en liefde was dit proefschrift er niet geweest. Het is fantastisch om met jou de afsluiting van allebei onze

PhD's te kunnen vieren door drie maanden op reis te gaan naar de mooiste, hoogste en bijzonderste plekken van de wereld.  Ik houd van je, en ik kijk uit naar alle mooie jaren die nog samen zullen komen!

# List of publications

## Included in this manuscript

International League Against Epilepsy Consortium on Complex Epilepsies. (2018). Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nature Communications*, *9*(1), 5269.

Leu C, **Stevelink R**, Smith AW, Goleva SB, Kanai M, Ferguson L, Campbell C, Kamatani Y, Okada Y, Sisodiya SM, Cavalleri GL, Koeleman BPC, Lerche H, Jehi L, Davis LK, Najm IM, Palotie A, Daly MJ, Busch RM, Epi25 Consortium & Lal D. (2019). Polygenic burden in focal and generalized epilepsies. *Brain*, *142*(11), 3473–3481.

Mirza N, **Stevelink R**, Taweel B, Koeleman BPC, Marson AG & International League Against Epilepsy Consortium on Complex Epilepsies. (2021). Using common genetic variants to find drugs for common epilepsies. *Brain Communications*, *3*(4), fcab287.

**Stevelink R**, Koeleman BPC, Sander JW, Jansen FE & Braun KPJ. (2019). Refractory juvenile myoclonic epilepsy: a meta-analysis of prevalence and risk factors. *European Journal of Neurology; 26*(6), 856–864.

**Stevelink R**, Luykx JJ, Lin BD, Leu C, Lal D, Smith AW, Schijven D, Carpay JA, Rademaker K, Rodrigues Baldez RA, Devinsky O, Braun KPJ, Jansen FE, Smit DJA, Koeleman BPC, International League Against Epilepsy Consortium on Complex Epilepsies & Epi25 Collaborative. (2021). Shared genetic basis between genetic generalized epilepsy and background electroencephalographic oscillations. *Epilepsia*, *62*(7), 1518–1527.

**Stevelink R**, Pangilinan F, Jansen FE, Braun KPJ, International League Against Epilepsy Consortium on Complex Epilepsies, Molloy AM, Brody LC & Koeleman BPC. (2019). Assessing the genetic association between vitamin B6 metabolism and genetic generalized epilepsy. *Molecular Genetics and Metabolism Reports*, *21*, 100518.

**Stevelink R**, Sanders MW, Tuinman MP, Brilstra EH, Koeleman BPC, Jansen FE, & Braun KPJ. (2018). Epilepsy surgery for patients with genetic refractory epilepsy: a systematic review. *Epileptic Disorders; 20*(2), 99–115.

International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide meta-analysis of over of 29,000 people with epilepsy reveals 26 loci and subtype-specific genetic architecture. *Under review.*

**Stevelink R**, Al-Toma D, Jansen FE, Lamberink HJ, Asadi-Pooya AA, Farazdaghi M, Cação G, Jayalakshmi S, Patil A, Özkara C, Aydın S, Gesche J, Beier CP, Stephen LJ, Brodie MJ, Unnithan G, Radhakrishnan A, Höfler J, Trinka E, Krause R, EpiPGX Consortium, Cerulli Irelli E, Di Bonaventura C, Szaflarski JP, Hernández-Vanegas LE, Moya-Alfaro ML, Zhang Y, Zhou D, Pietrafusa N, Specchio N, Japaridze G, Beniczky S, Janmohamed M, Kwan P, Syvertsen M, Selmer KK, Vorderwülbecke BJ, Holtkamp M, Viswanathan LG, Sinha S, Baykan B, Altindag E, von Podewils F, Schulz J, Seneviratne U, Viloria-Alebesque A, Karakis I, D'Souza WJ, Sander JW, Koeleman BPC, Otte WM & Braun KPJ (2022). Individualised prediction of drug resistance and seizure recurrence after medication withdrawal in people with juvenile myoclonic epilepsy: a systematic review and individual participant data meta-analysis. *Under review.*

**Stevelink R**, Koeleman BPC, Sisodiya SM & International League against Epilepsy Consortium on Complex epilepsies. (2022). Distinct genetic basis of common epilepsies and structural MRI measures. *Manuscript under review.*

## Other publications

**Stevelink R** & Kok G. (2018). Young scientists aim to prioritize patients. *Nature*, *558*(7711), 519.

Schijven D, **Stevelink R**, McCormack M, van Rheenen W, Luykx J J, Koeleman BPC, Veldink JH, Project MinE ALS GWAS Consortium & International League Against Epilepsy Consortium on Complex Epilepsies. (2020). Analysis of shared common genetic risk between amyotrophic lateral sclerosis and epilepsy. *Neurobiology of Aging*, *92*, 153.e1-153.e5.

Niestroj L-M, Perez-Palma E, Howrigan DP, Zhou Y, Cheng F, Saarentaus E, Nürnberg P, **Stevelink R**, Daly MJ, Palotie A, Lal D & Epi25 Collaborative. (2020). Epilepsy subtype-specific copy number burden observed in a genome-wide study of 17 458 subjects. *Brain*, *143*(7), 2106–2118.

Dijkstra S, Kok G, Ledford JG, Sandalova E & **Stevelink R**. (2018). Possibilities and Pitfalls of Social Media for Translational Medicine. *Frontiers in Medicine*, *5*, 345.

**Stevelink R**, Abramovic L, Verkooijen S, Begemann MJH, Sommer IEC, Boks MP, Mandl RC W, van Haren NEM & Vinkers CH. (2018). Childhood abuse and white matter integrity in bipolar disorder patients and healthy controls. *European Neuropsychopharmacology*, *28*(7), 807–817.

Verkooijen S, **Stevelink R**, Abramovic L, Vinkers CH, Ophoff RA, Kahn RS, Boks MPM & van Haren NEM. (2017). The association of sleep and physical activity with integrity of white matter microstructure in bipolar disorder patients and healthy controls. *Psychiatry Research: Neuroimaging*, *262*, 71–80.

Schene IF, Joore IP, Baijens JHL, **Stevelink R**, Kok G, Shehata S, Ilcken EF, Nieuwenhuis ECM, Bolhuis DP, van Rees RCM, Spelier SA, van der Doef HPJ, Beekman JM, Houwen RHJ, Nieuwenhuis EES & Fuchs SA. (2022). Mutation-specific reporter for optimization and enrichment of prime editing. *Nature Communications. 13(1)*, 1028.

Gudberg C, **Stevelink R**, Douaud G, Wulff K, Lazari A, Fleming MK & Johansen-Berg H. (2022) Individual differences in slow wave sleep architecture relate to variation in white matter microstructure across adulthood. *Frontiers in Aging Neuroscience.* In press.

# About the author

Remi Stevelink was born on 6 October 1990 in Eelde, a village in the north of the Netherlands, where he was raised by his loving parents together with his brother Elwin and sister Lisa.

After finishing high school in 2008, he spent a year to travel around the world, visiting 12 countries in Asia, North- and South America. Afterwards, he applied for medicine but did not win the lottery, partly due to poor grades in high school. After a week of studying biology, he got admitted to biomedical sciences. During his bachelor studies he spent a semester as exchange student in Hong Kong.

Unsure what to do next, he took another gap year to travel in the Middle-East and Africa, and perform internships in neuroscience (Lausanne, Switzerland) and medicine (Nakuru, Kenya). He realized he was interested in medicine as well as neuroscience and decided to apply for both. Fortunately, he was able to combine these fields by first finishing a neuroscience master's programme in Oxford, after which he returned to Utrecht in 2014 to start a graduate medicine programme.

After a brief stint in psychiatry research, he combined his medicine studies with research in paediatric neurology and genetics, supervised by Bobby Koeleman, Floor Jansen and Kees Braun, which eventually grew into the PhD research that resulted in this thesis. He graduated from medical school in 2018, after which he started a neurology residency programme in Utrecht. He aims to become a paediatric neurologist, and to continue to combine his passions of clinical work and research.

He lives in Utrecht with his girlfriend Juliëtte and their cat Chouffe. In his free time he loves travelling to remote countries, hiking, road cycling and enjoying good beers with friends.