RESEARCH ARTICLE

Integrative Deep Learning of Genomic and Clinical Data for Predicting Treatment Response in Newly Diagnosed Epilepsy

Wei Feng,[1,2,3,*] Duong Nhu,[2,4,5,*] Alison Anderson,[6,*] Daniel Thom,[6,7,8] Sarah N. Barnard,[6,7] Robert Zeibich,[6] Emma Foster,[6,7] Mark Howard,[9] Susannah T. Bellows,[10] Rosemary Burgess,[10] Samuel F. Berkovic,[10,11] Terence J. O'Brien,[6,7] Zhibin Chen,[6] Jacqueline French,[12] Patrick Kwan,[6,7,8,†] and Zongyuan Ge[1,2,3,4,5,†]

Correspondence
Dr. Kwan
patrick.kwan@monash.edu

Neurology® 2025;105:e214315. doi:10.1212/WNL.0000000000214315

Abstract

Background and Objectives

Epilepsy is a common neurologic disorder. Although antiseizure medications (ASMs) are the first-line treatment, identifying the most effective ASM for each individual remains a trial-and-error process. Genetic variation may influence treatment response. We aimed to develop and validate a multimodal deep learning model that integrates clinical and genomic features to predict response to the initial ASM in people with newly diagnosed epilepsy.

Methods

We used data from individuals with newly diagnosed epilepsy in Australia as the development cohort and participants from the Human Epilepsy Project 1 (recruited in the United States, Europe, and Australia) as the external validation cohort. All participants initiated ASM treatment and were followed prospectively for at least 1 year. We included 16 clinical factors and constructed 4 genomic feature types related to epilepsy and ASM pharmacogenomics, with and without functional impact annotations. We evaluated various machine learning architectures and multimodal fusion strategies to predict seizure freedom while taking the initial ASM at 1 year.

Results

In the development cohort (n = 286, median age 39 years, 47.2% seizure free), combining clinical and genomic features in our proposed multimodal deep learning model improved predictive performance. The highest area under the receiver operating characteristic curve (AUC) of 0.74 (95% CI 0.70–0.78) was achieved using clinical factors and genomic variants affecting transcription factor binding, significantly outperforming the clinical-only model (AUC 0.67, 95% CI 0.62–0.72; $p < 0.05$). In the external validation cohort (n = 219, median age 31 years, 20.5% seizure free), the same feature combination achieved an AUC of 0.69 (95% CI 0.67–0.71), higher than the clinical-only model (AUC 0.62, 95% CI 0.60–0.64; $p < 0.05$). Applying this model to the development cohort, if all participants took the highest ranked ASMs, the mean predicted seizure-free probability would be 68.05% (95% CI 65.79%–70.35%) compared with the observed seizure-free rate of 47.2% (95% CI 41.3%–53.2%).

Discussion

Integrating genomic data with clinical features enhances the ability of deep learning models in predicting ASM response in newly diagnosed epilepsy. This approach may support personalized treatment selection and improve clinical outcomes.

RELATED ARTICLE

Editorial
Deep Learning for the Prediction of Treatment Responses in Individuals With Epilepsy: Can We Get to Seizure Freedom Faster?
Page e214421

MORE ONLINE

Supplementary Material

*These authors contributed equally to this work.

†These authors contributed equally to this work as co-senior authors.

[1]Faculty of Engineering, Monash University, Melbourne, Australia; [2]Monash Medical AI, Faculty of Information Technology, Monash University, Melbourne, Australia; [3]Airdoc-Monash Research Lab, Monash University, Melbourne, Australia; [4]Faculty of Information Technology, Monash University, Melbourne, Australia; [5]AIM for Health Lab, Monash University, Melbourne, Australia; [6]Department of Neuroscience, School of Translational Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia; [7]Department of Neurology, Alfred Health, Melbourne, Australia; [8]Department of Neurology, Royal Melbourne Hospital, Australia; [9]Department of Philosophy, School of Philosophical, Historical, and Indigenous Studies, Faculty of Arts, Monash University, Melbourne, Australia; [10]Epilepsy Research Centre, Department of Medicine, Austin Health, The University of Melbourne, Australia; [11]Bladin-Berkovic Comprehensive Epilepsy Program, Department of Neurology, Austin Health, Melbourne, Australia; and [12]New York University Grossman School of Medicine and NYU Langone Health, Comprehensive Epilepsy Center, NY.

Glossary

ASM = antiseizure medication; **AUC** = area under the receiver operating characteristic curve; **CV** = cross-validation; **eQTL** = expression quantitative trait loci; **GBDT** = Gradient Boosting Decision Tree; **GOI** = gene of interest; **GT** = genotypes; **HEP1** = Human Epilepsy Project 1; **HOX** = homeobox; **MELB** = Melbourne cohort; **MFB** = multimodal factorized bilinear; **MLP** = multilayer perceptron; **SHAP** = Shapley Additive Explanation; **SNP** = single-nucleotide polymorphism; **TF** = transcription factor.

Introduction

One in 26 people will develop epilepsy during their lifetime.[1] Compared with the general population, people with epilepsy have increased medical and psychiatric comorbidities, increased risk of premature death,[2] reduced quality of life,[3] and reduced productivity, placing a substantial economic cost on them and society.[1] Recurrent seizures can be suppressed by treatment with antiseizure medications (ASMs) in a proportion of patients. However, there is high variability in response to ASMs, ranging from complete seizure freedom to ongoing seizures and/or the development of intolerable adverse effects that result in early ASM discontinuation(s) and switching. The challenge of treatment selection is compounded by the introduction of more than 15 new ASMs to clinical practice in the past 2 decades, most of which have shown similar efficacy when compared on a group basis.[4]

It has been proposed that a more "personalized" approach in ASM selection that takes into account personal characteristics may improve treatment outcomes for people with epilepsy.[5] Previous research has identified a number of clinical predictors of response to ASMs, including type of epilepsy, number of pretreatment seizures, and presence of psychiatric comorbidities.[6] Based on the aforementioned factors, a machine learning model was recently developed that demonstrated the feasibility of personalized prediction of response to the first ASM in adults with newly diagnosed epilepsy.[7]

In addition to clinical factors, genomic variants are known to influence treatment response through their pharmacokinetic and pharmacodynamic effects.[8] The incorporation of DNA sequencing data into machine learning models is technically challenging because of its high dimensionality but has been performed. A recent study demonstrated that machine learning models can learn patterns from the combination of clinical and genomic factors to predict the outcome of adjunctive brivaracetam in individuals with drug-resistant epilepsy enrolled in regulatory clinical trials.[8] This study explored the predictive value of structural variants that affect the molecular target of brivaracetam (synaptic vesicle 2A), as well as genes related to the development of epilepsy, providing proof-of-concept of this approach. However, the abovementioned methods are still limited to studying a single ASM, brivaracetam, and research on other ASMs is still insufficient. In addition, existing studies are based only on limited genomic data and fail to fully explore the potential impact of different genomic features on treatment response. It is important to note that most existing methods fail to fully harness the significant potential of deep learning models in capturing complex data patterns and performing cross-modal feature fusion, thus limiting their application in efficient and accurate diagnosis and personalized treatment plans. Beyond technical challenges, there are also ethical concerns, such as privacy, consent, distributive justice/fairness (including generalizability), data bias, and data discrimination, which need to be considered in the use of genomic data in machine learning models.[9,10]

Our ultimate research aim is to develop a machine learning model that may be applied to increase the likelihood of seizure freedom by assisting clinicians to select the ASM predicted to be most effective on a patient-by-patient basis. To develop the model, we used the actual clinical outcome of each patient's first prescribed ASM as the ground truth. We designed a multimodal deep learning architecture that captured both seizure control and drug tolerance. To address the high dimensionality of genomic data, we constructed genomic-specific features. We assessed how clinical data and genetic composition uniquely contribute to predicting drug response. Model performance was then compared across unimodal inputs (clinical or genomic features alone) and multimodal inputs (combined clinical and genomic features) to evaluate the added value of integrating multiple data types. We demonstrated that our best multimodal model achieved good predictive performance in an independent external cohort. Finally, we illustrated how the model might be applied clinically in the future.

Methods

Standard Protocol Approvals, Registrations, and Patient Consents

The development cohort (Melbourne cohort, MELB) was approved by the Melbourne Health Human Research Ethics Committee (Ethics Approval Number: HREC 2002.232). All participants provided written informed consent. The external validation cohort from the multicenter Human Epilepsy Project 1 (HEP1) was approved centrally by the New York University Institutional Review Board and locally by each participating site's institutional research ethics board. Participants in HEP1 were recruited from 34 major tertiary epilepsy centers across the United States, Australia, Canada, and Europe within 4 months of initiating ASM treatment for

newly diagnosed focal epilepsy. The handling of sensitive patient information, including genomic data, was conducted in strict accordance with ethical guidelines and recognized key bioethics principles such as justice and nonmaleficence. Data were stored on secure, access-controlled servers, and only authorized members of the research team had access. The types of seizures and epileptic syndromes were recorded in each study and classified according to the International League Against Epilepsy criteria.[11] Full details of cohort characteristics and study procedures are provided in eMethod 1.

Study Cohorts

Development Cohort (MELB)

Individuals with newly diagnosed epilepsy were prospectively recruited at 2 national hospitals in Melbourne, Australia (2003–2016). The study setting of the development cohort has been previously described.[7] Eligible participants were at least 9 years old, had initiated ASMs within 3 months before recruitment, and were followed for at least 1 year. Those with previous epilepsy or untreated with ASMs during the study period were excluded.

External Validation Cohort (HEP1)

The external validation cohort comprised participants from the HEP1,[12] a multinational and multicenter prospective study (2012–2020) focused on newly diagnosed focal epilepsy. Study eligibility was determined according to the inclusion/exclusion criteria as outlined in previous publications.[13] Principal component analysis was used to remove genetic outliers based on ancestry.

Outcome Assessment

The participants of the development cohort were prospectively followed up at 3 months and 1 year after enrollment. At each follow-up, the responses to therapy were recorded, including seizure occurrence and emergence of adverse events. Participants of the external validation cohort had standard annual follow-up visits for up to 6 years after enrollment, which usually coincided with a clinical visit. Seizure and ASM information was prospectively collected and entered into the participant's personal electronic seizure diary. Data from the diaries were verified with participants on annual visits and cross-referenced with their medical records.

For this analysis, the outcome was based on the achievement of seizure freedom during the first year of commencing treatment. Seizure freedom was defined as no seizure within 12 months of treatment initiation and continued use of the first ASM monotherapy at 12 months after treatment initiation. Treatment failure was considered if ASM therapy was changed (switched/added) for any reason within the first 12 months.

Clinical Factors

We included the same clinical factors used in a previous study that developed a deep learning model to predict treatment response to the first ASM in a different cohort of people with newly diagnosed epilepsy.[7] These factors were sex, history of cerebral infection, significant head trauma, febrile seizure, cerebral hypoxic injury, substance abuse, alcohol abuse, epilepsy in first-degree relative, cerebrovascular disease, intellectual disability, psychiatric disorder, number of pretreatment seizures, type of epilepsy, EEG findings, MRI findings, and the first prescribed ASM.

Genotypes

Participants in both cohorts had been previously genotyped using single-nucleotide polymorphism (SNP) arrays, and these data were imputed against the same reference panel and the same quality filtering was applied (eMethod 2 for details).

Construction of Genomic Features

We constructed 4 genomic features (Table 1). Each of these comprised an input matrix, whereby the rows were genes, the columns were sample IDs, and the cell values represented the burden of genetic variation (eMethods 2 provides full details). In brief, we generated 2 types of genomic features (which were not mutually exclusive), one comprising genotype information only (Table 1, GT Feature) and one that included scores representing the functional consequence of the alternative genotype (remaining 3 Features). We applied a targeted approach based on genes associated with epilepsy or pharmacogenomics of ASMs (in eTables 1–3), hereafter referred to as genes of interest (GOIs). To construct the GT Feature, the genotypes for variants within GOIs for each individual were encoded as 0 for noncarriers, 1 for heterozygous carriers, and 2 for homozygous carriers and aggregated at the gene level to provide an input matrix with 1,023 rows. To capture the impact of genetic variation on gene expression, we used normalized effect size scores from the Genotype-Tissue Expression project,[14] commonly referred to as expression quantitative trait loci (eQTL) data, specific to brain tissue. We generated 2 features whereby the eQTL scores for each variant were multiplied by an individual's genotype (0, 1, 2) at the variant locus and aggregated at the gene level. One feature represented the burden of variants that fall within a GOI but affect the expression of any gene (eQTL-wGOI Feature), and the other comprised variants within other genes that are predicted to affect the expression of a GOI (eQTL-iGOI Feature). For the remaining feature, we focused on genes that encode transcription factor (TF) proteins with the aim to capture the impact of genetic variation on gene regulation genome-wide (TF Feature).[15] We used the published DeepBind tool[16] to generate scores representing the likelihood that a variant will affect the binding affinity of 1 or more TFs. eFigure 1 illustrates the distribution and cutoff selection of DeepBind scores across TFs for generating TF Features. Genotypes were multiplied by the DeepBind score to generate a burden score for each variant-disrupted binding site, and scores were aggregated at the TF level. The resulting TF Feature comprised all TFs for which binding affinity had been predicted to be modified by at least 1 variant in at least 1 individual.

Table 1 Genomic Features

Genomic feature	Description	No. of genes implicated
GT	Encoded GT for variants within GOI	1,023
eQTL-wGOI	GT × eQTL scores for variants within a GOI that affect a GOI or any other genes	3,304
eQTL-iGOI	GT × eQTL scores for variants that affect a GOI	708
TF	GT × score for impact on TF binding	422

Abbreviations: GT = genotype; GOI = gene of interest; eQTL = expression quantitative trait loci; TF = transcription factor; wGOI = within GOI.

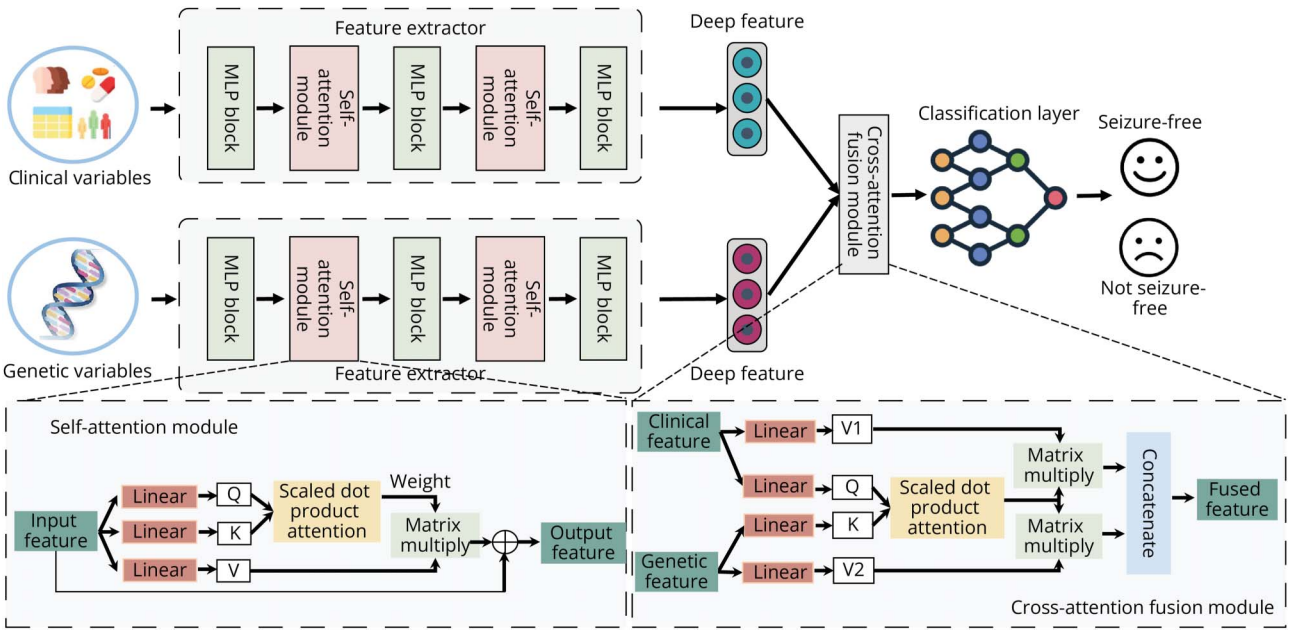
Model Development

Figure 1 illustrates the overall framework of our multimodal model. Two separate feature extractors/encoders were used to extract corresponding deep feature representations from clinical factors and genomic inputs. Each feature extractor consisted of 3 multilayer perceptron (MLP) blocks, each containing a linear layer, a batch normalization layer, a rectifier linear unit layer, and a dropout layer. To explore the relationships within each modality, we applied a self-attention module that exploited the associations across different feature variables within each modality based on the attention mechanism.[17] To further explore the correlation between the 2 modalities, we developed a cross-attention fusion module to integrate feature representations from both clinical factors and genomic features. After obtaining the fused feature representations, the classification layer, consisting of 3 linear layers and a softmax classifier, was used to derive the final prediction. Additional detailed steps for the development a Audio Volume Mutend optimization of multimodal deep learning models are presented in eMethod 3, and the detailed model parameter settings are provided in eTable 4.

Model Performance Evaluation

Consistent with our previous study,[7] we used the area under the receiver operating characteristic curve (AUC) to evaluate the performance of the models in predicting whether a participant would be seizure free while taking the first prescribed ASM during the first year of treatment. Model evaluation was performed in 3 experiments. First, in the *unimodality analysis* experiment, we compared the predictive performance of 3 unimodal models, namely Gradient Boosting Decision Tree (GBDT),[8] MLP, and ours (without cross-attention, only self-attention) for individual features (clinical alone and the individual genomic features). GBDT is a traditional machine learning model while MLP and our model are deep learning models. Second, in the *multimodality analysis* experiment, we used different combinations of clinical factors and genomic features and compared the performance of our multimodal method with that of 3 other multimodal methods: Mutan,[18] Block,[19] and multimodal factorized bilinear (MFB).[20] Third, in the *generalizability assessment* experiment, the model showing the best performance in the development cohort was tested in the external validation cohort. Moreover, because

Figure 1 Overall Framework of Our Multimodal Model



the HEP1 cohort only includes patients with focal epilepsy, we retrained the model using only participants with focal epilepsy in the MELB cohort to avoid any potential bias and evaluated the model on HEP1 accordingly. In addition, Shapley Additive Explanation (SHAP)[21] on the development cohort was used to help interpret our model.

Statistical Analysis

We applied 5-fold cross-validation (CV) in which each data set was divided into 5 folds and each fold was iteratively left out for testing, while the rest were used to train the model. We applied the bias-corrected and accelerated bootstrap CI method with 1,000 repetitions to estimate the mean and 95% CI for the AUC in each fold of CV. Pooled estimates were obtained by random-effects meta-analysis with the Sidik-Jonkman estimator and robust variance to account for within-fold and between-fold variability and fold dependence.[22] Model performance comparisons were subsequently conducted using random-effects meta-regression with the same specifications and the Benjamini-Krieger-Yekutieli procedure to control the false discovery rate at 5% in pairwise tests. We also analyzed the association of the top contributing genomic features with ASM response by calculating the standardized mean difference (Hedges g) between the seizure-free and non–seizure-free groups. Statistical significance was set at $p < 0.05$ unless otherwise specified. All statistical analyses were conducted using Python version 3.10 (Python Software Foundation) with Scipy 1.15.3 and R version 4.4.2 (R Core Team).

To illustrate how our model might be used to support treatment selection in the future, we used it to derive the predicted probability of seizure freedom associated with each ASM for each patient in the development cohort and ranked the ASMs accordingly. We then calculated the highest predicted probability of seizure freedom for the whole cohort based on the assumption that each patient would receive the ASM with the highest ranked predicted probability of seizure freedom. This was indirectly compared with the actual seizure-free rate observed in the cohort based on the prescribed ASMs. In addition, we assessed the distribution of predicted probabilities across the cohort using a hypothetical cutoff from the best-performing folds to categorize patients into 3 different groups. We analyzed feature associations using χ^2 tests for categorical clinical factors and independent t tests for TF Feature scores of the top 20 SHAP-derived TFs.

Data Availability

Anonymized data not published within this article will be made available by request from any qualified investigator.

Results

Clinical and Genetic Characteristics

The final development cohort included 286 participants, and the external validation cohort had 254 participants. Most of the participants in both cohorts were White European (MELB = 90% [259/286], HEP1 = 79% [202/254]). Anchoring of the genotypes to data from the 1000 Genomes Project[23] identified individuals in the external validation cohort, with ethnicities not represented in the development cohort (Black American, n = 24; Black Asian, n = 1; White Asian, n = 1) or that clustered with this group (unknown, n = 6; White, n = 3). These individuals (35/254, 14%) were excluded from further analysis (eFigure 2; eTables 5 and 6), leaving a total of 219 for external validation (92% [202/219] were White).

The clinical characteristics used for model development and the ASMs prescribed for the participants are summarized in Table 2. In accordance with their eligibility criteria, the 2 cohorts differed in the proportions of individuals with a history of substance or alcohol abuse and epilepsy types. A higher proportion of female individuals and psychiatric disorders were observed in the external validation cohort. Note that the HEP1 cohort included patients aged ≥12 and ≤60 years at the time of seizure diagnosis, which differs in age distribution from the development cohort. Other clinical factors were comparable between the 2 cohorts. Carbamazepine and valproate were most commonly used in the development cohort, whereas most individuals were prescribed levetiracetam in the external validation cohort. In the development cohort, 135 individuals (47.2%) were seizure free during the first year after treatment, whereas in the external validation cohort, 45 individuals (20.5%) were seizure free during the first year after treatment.

The number of qualifying variants was greater for the validation cohort (n = 13,015,830) than the development cohort (n = 9,069,919), with 7,005,382 variants being present in 1 or more individuals from both cohorts. Among GOIs, there was a 78% overlap in variants involving 993 genes, and for 14 of these, variants were present in either cohort but not both (eTable 3). Genetic analysis confirmed that there was no overlap in individuals between the 2 cohorts.

Experiment 1: Unimodality Analysis

As shown in Figure 2A, when using a single data modality, our model achieved superior classification performance with clinical features (AUC 0.67, 95% CI 0.62–0.72) compared with any genomic feature (all corrected p values <0.001). Among the genomic modalities, eQTL-wGOI and TF yielded comparable AUCs of 0.62 (95% CI 0.60–0.64) and 0.61 (95% CI 0.59–0.63), respectively.

Furthermore, in the unimodal setting, our method and MLP consistently outperformed the traditional GBDT model (all corrected p values <0.05), highlighting the superior learning capacity of deep learning models. Across all cases, the mean AUC of the MLP model was 0.57, compared with 0.51 for GBDT. Finally, our approach further improved the AUC from 0.57 to 0.59 by incorporating a self-attention mechanism to integrate features across all modalities (all corrected p values <0.05).

Table 2 Clinical Characteristics of Participants and the Antiseizure Medications Prescribed in the Development Cohort and External Validation Cohort

Characteristic	Cohort, no. of participants (%)	
	Development cohort (N = 286)	External validation cohort (N = 219)
Sex		
Male	152 (53.1)	84 (38.3)
Female	134 (46.9)	135 (61.7)
Age at treatment initiation, y, median (IQR)	39 (24–55)	31 (20–41)
History of febrile convulsions		
Yes	11 (3.8)	6 (2.7)
No	271 (94.7)	150 (68.5)
N/A	4 (1.4)	63 (28.8)
History of CNS infection in childhood		
Yes	1 (0.3)	3 (1.3)
No	281 (98.3)	199 (90.9)
N/A	4 (1.4)	17 (7.8)
History of significant head trauma		
Yes	48 (16.8)	14 (6.4)
No	235 (82.2)	189 (86.3)
N/A	3 (1.0)	16 (7.3)
History of cerebral hypoxic injury		
Yes	3 (1.0)	1 (0.4)
No	279 (97.6)	202 (92.2)
N/A	4 (1.4)	16 (7.4)
History of substance abuse		
Yes	108 (37.8)	0
No	17 (5.9)	219 (100.0)
N/A	161 (56.3)	0
History of alcohol abuse		
Yes	103 (36.0)	0
No	20 (7.0)	219 (100.0)
N/A	163 (57.0)	0
History of epilepsy in first-degree relative		
Yes	83 (29.0)	33 (15.1)
No	201 (70.3)	179 (81.7)
N/A	2 (0.7)	7 (3.2)
History of cerebrovascular disease		
Yes	20 (7.0)	5 (2.3)
No	263 (92.0)	197 (90.0)
N/A	3 (1.0)	17 (7.7)

Continued

Table 2 Clinical Characteristics of Participants and the Antiseizure Medications Prescribed in the Development Cohort and External Validation Cohort *(continued)*

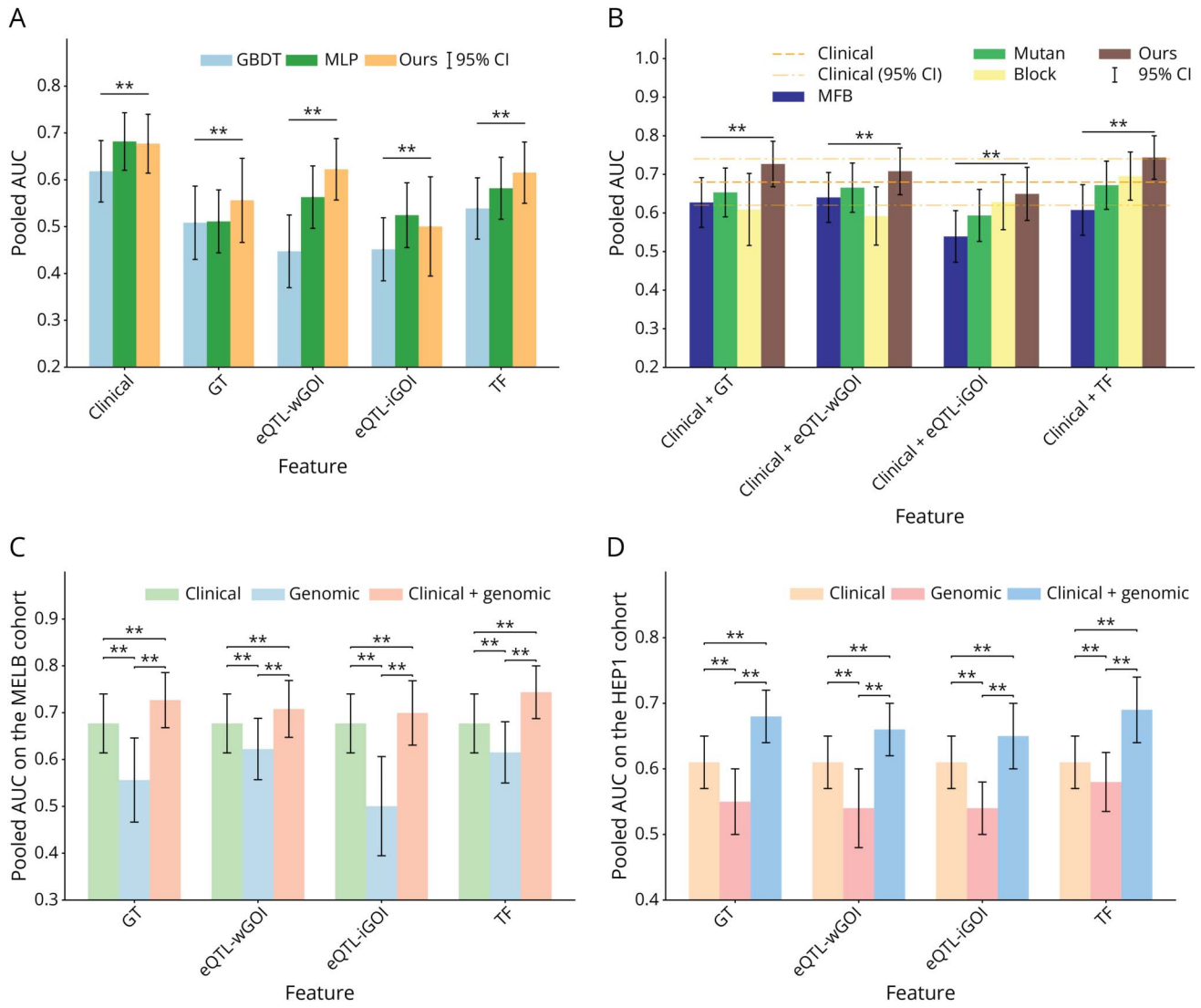
Characteristic	Cohort, no. of participants (%)	
	Development cohort (N = 286)	External validation cohort (N = 219)
History of intellectual disability		
Yes	6 (2.1)	22 (10.0)
No	278 (97.2)	197 (90.0)
N/A	2 (0.7)	0
History of psychiatric disorder		
Yes	23 (8.0)	65 (29.7)
No	105 (36.7)	153 (69.9)
N/A	158 (55.3)	1 (0.4)
No. of pretreatment seizures		
≤5	157 (54.9)	103 (47.0)
>5	128 (44.8)	115 (52.5)
N/A	1 (0.3)	1 (0.5)
Type of epilepsy		
Focal	223 (78.0)	219 (100.0)
Generalized	61 (21.3)	0
Unknown	2 (0.7)	0
EEG findings		
Epileptiform abnormality	97 (33.9)	94 (42.9)
Nonepileptiform abnormality	34 (11.9)	37 (16.9)
Normal	155 (54.2)	86 (39.3)
N/A	0	2 (0.9)
MRI findings		
Epileptogenic abnormality	49 (17.1)	40 (18.2)
Nonepileptogenic abnormality	54 (18.9)	46 (21.0)
Normal	174 (60.8)	121 (55.3)
N/A	9 (3.2)	12 (5.5)
First prescribed antiseizure medication		
Levetiracetam	6 (2.1)	151 (68.9)
Carbamazepine	137 (47.9)	5 (2.3)
Valproate	99 (34.6)	1 (0.5)
Lamotrigine	13 (4.5)	52 (23.7)
Phenytoin	31 (10.9)	10 (4.6)

Abbreviations: IQR = interquartile range; N/A = not available.

Experiment 2: Multimodality Analysis

In the multimodal setting, we integrated clinical variables and genomic features as model inputs. As illustrated in Figure 2B, our multimodal model using a cross-attention fusion strategy achieved the highest mean AUC of 0.70, followed by the MFB model (mean AUC 0.63). As shown in Figure 2C, the

Figure 2 Performance of Different Architectures and Combinations of Features



"Clinical" means the model uses clinical factors only; "GT" represents genotypes-only features; "eQTL-wGOI" refers to eQTLs located within genes of interest (GOI); "eQTL-iGOI" denotes eQTLs that affect GOI expression; and "TF" indicates transcription factor features affecting GOIs. (A) Pooled AUC scores of 5-fold CV of different models trained on genomic data or clinical data in the development cohort (MELB). (B) Pooled AUC scores of the 5-fold CV with different multimodal methods in the development cohort (MELB). (C) Performance comparison between our multimodal model and unimodal models in the development cohort (MELB). (D) Performance comparison between our multimodal model between unimodal models in the external validation cohort (HEP1). $^{**}p < 0.01$.

multimodal models consistently outperformed the unimodal counterparts across all feature combinations (all corrected p values <0.05). In particular, the TF Feature, which represented the consequences of TF binding perturbation, demonstrated notable predictive power (AUC 0.74, 95% CI 0.70–0.78). A detailed quantitative performance report is provided in eTable 7.

Experiment 3: Generalizability Assessment

In the external validation experiment (a), where the trained model on the whole development cohort was tested on the external cohort, the multimodal model consistently outperformed both the unimodal clinical and genomic models (Figure 2D), demonstrating superior predictive ability and robustness (all corrected p values <0.001). Notably, combining clinical features with either the TF Feature (AUC 0.69, 95% CI 0.67–0.71) or the GT Feature (AUC 0.68, 95% CI 0.66–0.70) yielded better predictive performance. Similar to the development cohort, the model using only clinical features outperformed most models using genomic features alone (all corrected p values <0.05). A detailed quantitative performance report is provided in eTable 8. In the external validation experiment (b), where only patients with focal epilepsy in the development cohort were included in model training, the multimodal models maintained their superior performance compared with the unimodal models when tested in the external cohort (eTable 9; $p < 0.05$). This finding further supports the model's generalizability across different patient populations.

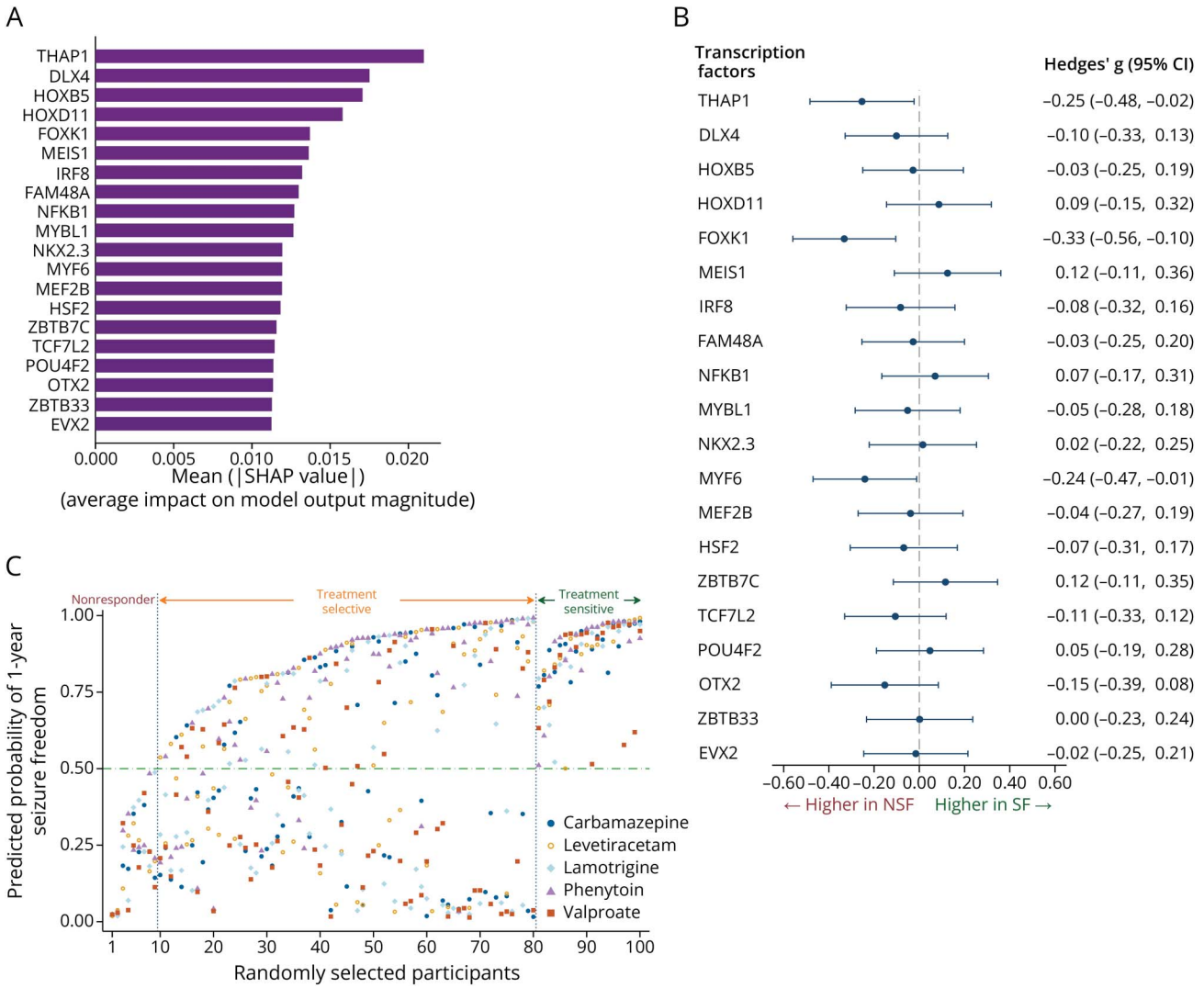
SHAP and Association Analysis of Contributing TFs

To specifically identify TFs with high predictive capacity, we restricted the SHAP analysis to the top 20 contributing TFs, as shown in Figure 3A. eFigure 3 shows the SHAP analysis displaying the top 20 contributing factors for the unimodal model using clinical factors only and for our final multimodal model (combined clinical factors and TF). The most important contributors include THAP1, DLX4, HOXB5, and HOXD11, among others. To demonstrate the association of these TFs with ASM response, Figure 3B shows the standardized mean difference of the TF scores (GT × score for impact on TF binding) between seizure-free and non–seizure-free groups. The analysis showed that variant-induced perturbation of THAP1 DNA binding was associated with the non–seizure-free outcome. Of these top 20 contributing TFs, 13 have been previously implicated in chemoresistance (eTable 10). Eight of the top 20 contributing TFs were homeobox (HOX) genes.

Potential Future Clinical Application

Applying the best-performing multimodal model (clinical + TF), Figure 3C shows the predicted probabilities of achieving seizure freedom for both prescribed and unprescribed ASMs across 100 randomly selected participants in the development cohort. Assuming that all participants would receive their

Figure 3 SHAP Analysis and Potential Clinical Application



(A) SHAP analysis of our final multimodal model in the development cohort (combined clinical factors and TF). The top 20 most contributing TFs are displayed. (B) Associations of the top 20 most contributing TFs with ASM response using the TF scores (GT × score for impact on TF binding). We calculated the standardized mean difference (Hedges *g*), as shown on the x-axis. The negative values mean the scores of the genes are higher in the non–seizure-free group, indicating the negative association with seizure freedom, and vice versa. NSF = not seizure-free; SF = seizure-free. (C) Predicted response to different antiseizure medications (ASMs) in 100 participants randomly selected from the development cohort. The horizontal green dash-dot line represents the hypothetical probability cutoff of 0.5 for seizure freedom. The vertical dotted lines divide the participants into 3 groups: "Non-responder" (all included ASMs have a predicted probability below the 0.5 cutoff for achieving seizure freedom), "Treatment Selective" (some ASMs have a predicted probability of ≥0.5 for achieving seizure freedom, but others are below the cutoff), and "Treatment Sensitive" (all ASMs predicted to have a probability of ≥0.5 for achieving seizure freedom).

highest ranked ASMs, the mean predicted seizure-free probability for the whole cohort was 68.05% (95% CI 65.79%–70.35%). In comparison, the seizure-free rate observed in the whole cohort based on the prescribed ASMs was only 47.2% (95% CI 41.3%–53.2%). Using a 0.5 probability threshold in Figure 3C, participants were classified into 3 groups: "Non-responder" (no ASMs ≥0.5), "Treatment Selective" (a mix of ASMs above and below 0.5), and "Treatment Sensitive" (all ASMs ≥0.5). eTable 11 provides the number of patients by each clinical factor among the 3 groups in the development cohort. Among the clinical factors, only MRI findings, the top clinical contributor based on SHAP analysis (eFigure 3), differed significantly between the "Non-responder" and "Treatment Sensitive" groups, and between "Treatment Selective" and "Treatment Sensitive" groups (eTable 12), suggesting that epileptogenic abnormalities negatively affect ASM outcomes. Among the top 20 contributing TFs, the variant disruption of binding affinity scores differed significantly between Non-responder and Treatment Sensitive groups for THAP1 only ($p < 0.05$).

Discussion

Using a targeted approach, our study demonstrated that integrating structured genomic data with clinical factors improved machine learning performance in predicting the success of initial ASM treatment in epilepsy. Our multimodal model, especially when incorporating variants affecting TF binding, achieved higher AUC and was validated in an external cohort, regardless of whether it was trained on all patients or only those with focal epilepsy. This approach could potentially inform individualized ASM selection to improve seizure-free outcomes.

We observed that model performance was lower in the external cohort, likely due to differences in study design, seizure severity, and ASM prescription patterns between the 2 cohorts. The development cohort had higher seizure-free rates and more frequent use of older ASMs, whereas the external cohort, recruited later, showed a shift toward newer ASMs such as lamotrigine and levetiracetam. Despite these differences, the multimodal model consistently outperformed the clinical-only model, confirming the added value of genomic data.

We also highlight the importance of choosing an appropriate modality fusion strategy because genomic data are often sparse and high-dimensional. In addition, different modalities may have large distributional differences between them, and the choice of modality fusion strategy is critically important. While Mutan and Block use multilinear tensor decomposition, MFB uses a 1-dimensional moving pooling window to reduce feature size. These methods assume linearity in features, which does not hold for the complex nature of genomic data. Unlike other fusion methods, our model uses self-attention and cross-modality attention to adaptively integrate heterogeneous data. This strategy led to better performance and improved interpretability, supporting its utility in integrating complex multimodal biomedical data.

For genomic features, functional information pertaining to the disruption of TF binding affinity (TF Feature) and the GT Feature gave the best performance in both cohorts. TFs have been reported to play a role in drug resistance in other conditions. For example, the TF nuclear factor kappa B alters the expression of multidrug resistance protein 1, contributing to drug resistance in cancer cells,[24] and c-Fos regulation of *NANOG* confers fluorouracil resistance.[25] Unlike the GT Feature and eQTL-based Features eQTL-wGOI and eQTL-iGOI, variants in the TF Feature were selected based on their impact on the genome-wide binding of 1 or more of 515 TFs, and thus not restricted to genes associated with epilepsy or epilepsy pharmacogenomics. Indeed, 8 of the top 20 contributing TFs were HOX genes. The HOX genes are regulated by hormones and implicated in resistance to therapies that block hormone receptors,[26] and the expression of HOXB5 and other members of this protein family have been shown to predict drug resistance in gliomas.[26,27] The Feature based on genotypes only outperformed the 2 genomic features that incorporated eQTL functional scores. eQTLs that fall within GOIs (eQTL-wGOIs) provided slightly better model performance than those that affect the expression of GOIs (eQTL-iGOIs). These trends support the importance of variant selection methods and suggest that, within the context of drug response, variants influencing regulatory mechanisms may be more relevant than protein-coding variants.

Overall, incorporating genomic features significantly improved the predictive performance of the model across various settings. This finding suggests that interactions between genomic and clinical data are crucial, highlighting the value of multimodal integration strategies. While such approaches have been used in cancer drug response prediction using multiomics data (SNP, transcriptomics, DNA methylation) and convolutional models,[28] their application in epilepsy is constrained by the lack of omics data from the same samples. Nonetheless, methods such as graph convolutional networks and the inclusion of drug characteristics or information on functional roles of genes/TFs may further enhance model accuracy.

To illustrate the potential clinical application, we showed that if all the patients were administered the highest ranked ASMs by the model, their predicted mean probability of achieving seizure freedom (68%) would be substantially higher than the observed seizure-free rate (47.2%) in the development cohort. Although indirect, this comparison implies that the model may have the potential to assist clinicians in selecting the most effective treatment for individual patients. Arguably, the "Treatment Selective" group would benefit the most from this approach, as they could potentially be administered the higher ranked ASMs at treatment initiation and achieve seizure freedom sooner than random drug selection. A

randomized controlled trial to prospectively evaluate the effectiveness of using such a predictive model to assist in the selection of the initial ASM is currently underway.[29]

This study has several limitations. Despite including the largest cohort to date with paired clinical and genomic data for predicting ASM response, the sample size remains limited, and only 5-fold CV was used. Although k-fold CV can produce high-variance and dependent test errors,[30] the consistent performance across both development and external validation cohorts, together with the added value of combining clinical and genomic features, not only mitigates this limitation but also aligns with TRIPOD guidelines for developing a prediction model for diagnostic and prognostic purposes.[31] Only 5 ASMs were analyzed, which may limit applicability, although these represent widely prescribed first-line treatments.[32] Seizure freedom was defined as 12 months without seizures, which may undermine ASMs requiring longer titration. The model also used EEG and MRI findings,[33,34] which may not always be available at treatment initiation. Furthermore, both cohorts were predominantly White (MELB 90%, HEP1 92%), and 13.7% of external samples were excluded because of unmatched ethnicity, possibly affecting generalizability. A lack of diversity in genomic data, typically biased in favor of White ethnicities, is a recognized and ongoing challenge to the field of genomic research, including machine learning model development.[9] Our future work will incorporate more diverse populations and validate models across regions and ethnicities to improve fairness, robustness, and equity, especially for historically underserved groups in epilepsy care.[35]

In conclusion, we have developed and validated a multimodal deep learning model that integrates clinical factors with structured genomic information to improve the prediction of treatment outcomes in individuals with newly diagnosed epilepsy. By enabling individualized prediction of responses to various ASMs, our model holds promise as a clinical decision-support tool to assist in drug selection and optimize treatment strategies. This work represents a step toward more personalized and effective epilepsy care.

Acknowledgment

The authors acknowledge Hamada Hamid Altalib, DO, MPH (Yale University), who led the bioinformatics team of HEP1.

Author Contributions

W. Feng: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. D. Nhu: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. A. Anderson: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data. D. Thom: major role in the acquisition of data. S.N. Barnard: drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data. R. Zeibich: drafting/revision of the manuscript for content, including medical writing for content. E. Foster: drafting/revision of the manuscript for content, including medical writing for content. M. Howard: drafting/revision of the manuscript for content, including medical writing for content. S.T. Bellows: drafting/revision of the manuscript for content, including medical writing for content. R. Burgess: drafting/revision of the manuscript for content, including medical writing for content. S.F. Berkovic: drafting/revision of the manuscript for content, including medical writing for content. T.J. O'Brien: drafting/revision of the manuscript for content, including medical writing for content. Z. Chen: drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data. J. French: drafting/revision of the manuscript for content, including medical writing for content. P. Kwan: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. Z. Ge: drafting/revision of the manuscript for content, including medical writing for content.

Study Funding

This work was funded by the National Health and Medical Research Council, Airdoc Philanthropic Grant.

Disclosure

S.N. Barnard receives salary support from the Epilepsy Study Consortium and the NIH. E. Foster and/or her institution have received grants from Avant Research Foundation, Brain Foundation Australia, Clive and Vera Ramaciotti Foundation (Ramaciotti Health Investment Grant), GPCE, LivaNova (Australia and USA), Lundbeck (Australia), Monash Partners STAR Clinician Fellowship, Monash University Early Career Postdoctoral Fellowship, Royal Australasian College of Physicians Fellows Research Establishment Fellowship, and Sylvia and Charles Viertel Charitable Foundation Clinical Investigator Award, outside the submitted work. S.F. Berkovic and/or his institution have received research support or consultancy fees from Praxis Precision Medicines, Eisai, Jazz Pharmaceuticals, LivaNova, Seer Medical, and UCB Pharma. M. Howard has received research support from Eisai. T.-J. O'Brien's institution has received research support from Australian and US government research granting organizations (NHMRC, MRFF, DOD, National Institute of Neurological Disorders and Stroke) and research and/or consultancy fees from Eisai, Jazz Pharmaceuticals, LivaNova, ES Therapeutics, Kinosis Pharmaceuticals, Supernus, and UCB Pharma. Z. Chen has received research support from UCB Pharma. J.A. French receives salary support from the Epilepsy Foundation and from the Epilepsy Study Consortium for consulting work and/or attending scientific advisory boards for Acadia Pharmaceuticals, Access Industries, Acuta Capital Partners, AFASCI Inc., Agrithera Inc., Alterity Therapeutics Limited, Angelini Pharma S.p.A., Autifony Therapeutics Limited, Axonis Therapeutics, Baergic Bio, Beacon

Biosignals Inc., Biogen, Biohaven Pharmaceuticals, Bloom Science Inc., Bright Minds Biosciences Inc., Camp4 Therapeutics Corporation, Capsida Biotherapeutics, Cerebral Therapeutics, Cerecin Inc., Cerevel, Ceribell, Cognizance Biomarkers, Cowen and Company LLC, Crossject, EcoR1 Capital, Eisai, Encoded Therapeutics, Engrail, Epalex, EpiMinder, Epitel Inc., Equilibre BioPharmaceuticals, Genentech Inc., Grin Therapeutics, iQure Pharma Inc., IQVIA RDS Inc., Janssen Pharmaceutica, Jazz Pharmaceuticals, Korro Bio Inc., Leal Therapeutics Inc., Lipocine, LivaNova, Longboard Pharmaceuticals, Marinus, Modulight.bio, Neumirna Therapeutics, Neurelis, Neurocrine, NeuroPace Inc., NeuroPro Therapeutics, Neuroventis, Neurona Therapeutics, Ono Pharmaceutical Co., Otsuka Pharmaceutical Development, Ovid Therapeutics Inc., Praxis, PureTech LTY Inc., Rapport Therapeutics Inc., Receptor Holdings Inc., Rivervest Venture Partners, Sage Therapeutics Inc., SK Life Sciences, Stoke, Supernus, Takeda, Taysha Gene Therapies, Third Rock Ventures LLC, UCB Inc., Ventus Therapeutics, Vida Ventures Management, and Xenon. She has also received research support from the Epilepsy Study Consortium (funded by Eisai and UCB), the Epilepsy Study Consortium/Epilepsy Foundation (funded by UCB), the GW/FACES/One8 Foundation, and the National Institute of Neurological Disorders and Stroke. P. Kwan has received research support or consultancy fees from Angelini, Eisai, Jazz Pharmaceuticals, LivaNova, SK Life Science, and UCB Pharma. Z. Ge and his laboratory have received research support or consultancy fees from Optain Health, Eyetelligence, and NVIDIA AI Tech Centre. All other authors report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology*® April 2, 2025. Accepted in final form August 27, 2025. Submitted and externally peer reviewed. The handling editor was Associate Editor Emily Johnson, MD, MPH.

References

1. Foster E, Chen Z, Zomer E, et al. The costs of epilepsy in Australia: a productivity-based analysis. *Neurology*. 2020;95(24):e3221-e3231. doi:10.1212/WNL.0000000000010862
2. Neligan A, Bell GS, Johnson AL, Goodridge DM, Shorvon SD, Sander JW. The long-term risk of premature mortality in people with epilepsy. *Brain*. 2011;134(pt 2):388-395. doi:10.1093/brain/awq378
3. Strine TW, Kobau R, Chapman DP, Thurman DJ, Price P, Balluz LS. Psychological distress, comorbidities, and health behaviors among U.S. adults with seizures: results from the 2002 National Health Interview Survey. *Epilepsia*. 2005;46(7):1133-1139. doi:10.1111/j.1528-1167.2005.01605.x
4. Perucca E, Brodie MJ, Kwan P, Tomson T. 30 years of second-generation antiseizure medications: impact and future perspectives. *Lancet Neurol*. 2020;19(6):544-556. doi:10.1016/S1474-4422(20)30035-1
5. Chen Z, Rollo B, Antonic-Baker A, et al. New era of personalised epilepsy management. *BMJ*. 2020;371:m3658. doi:10.1136/bmj.m3658
6. Mohanraj R, Brodie MJ. Diagnosing refractory epilepsy: response to sequential treatment schedules. *Eur J Neurol*. 2006;13(3):277-282. doi:10.1111/j.1468-1331.2006.01215.x
7. Hakeem H, Feng W, Chen Z, et al. Development and validation of a deep learning model for predicting treatment response in patients with newly diagnosed epilepsy. *JAMA Neurol*. 2022;79(10):986-996. doi:10.1001/jamaneurol.2022.2514
8. de Jong J, Cutcutache I, Page M, et al. Towards realizing the vision of precision medicine: AI based prediction of clinical drug response. *Brain*. 2021;144(6):1738-1750. doi:10.1093/brain/awab108
9. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in healthcare. *Annu Rev Biomed Data Sci*. 2021;4(1):123-144. doi:10.1146/annurev-biodatasci-092820-114757
10. Sen SK, Green ED, Hutter CM, Craven M, Ideker T, Di Francesco V. Opportunities for basic, clinical, and bioethics research at the intersection of machine learning and genomics. *Cell Genom*. 2024;4(1):100466. doi:10.1016/j.xgen.2023.100466
11. 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. doi:10.1111/epi.13709
12. Accessed October 2, 2025. humanepilepsyproject.org
13. Pellinen J, Tafuro E, Yang A, et al. Focal nonmotor versus motor seizures: the impact on diagnostic delay in focal epilepsy. *Epilepsia*. 2020;61(12):2643-2652. doi:10.1111/epi.16707
14. Accessed October 2, 2025. gtexportal.org/home/
15. Tseng CC, Wong MC, Liao WT, et al. Genetic variants in transcription factor binding sites in humans: triggered by natural selection and triggers of diseases. *Int J Mol Sci*. 2021;22(8):4187. doi:10.3390/ijms22084187
16. Alipanahi B, Delong A, Weirauch MT, Frey BJ. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat Biotechnol*. 2015;33(8):831-838. doi:10.1038/nbt.3300
17. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Proceeding of the Conference on Neural Information Processing Systems*. 2017;30(1):5998-6008. doi:10.48550/arXiv.1706.03762
18. Ben-Younes H, Cadene R, Cord M, et al. Mutan: multimodal tucker fusion for visual question answering. *Proceedings of the IEEE International Conference on Computer Vision*. IEEE; 2017:2612-2620.
19. Ben-Younes H, Cadene R, Thome N, Cord M. Block: bilinear superdiagonal fusion for visual question answering and visual relationship detection. *Proc AAAI Conf Artif Intell*. 2019;33(1):8102-8109. doi:10.1609/aaai.v33i01.33018102
20. Yu Z, Yu J, Fan J, et al. Multi-modal factorized bilinear pooling with co-attention learning for visual question answering. *Proceedings of the IEEE International Conference on Computer Vision*. IEEE; 2017:1821-1830.
21. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Proceeding of the Conference on Neural Information Processing Systems*. 2017;30(1):4768-4777. doi:10.48550/arXiv.1705.07874
22. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2
23. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. *Nucleic Acids Res*. 2020;48(D1):D941-D947. doi:10.1093/nar/gkz836
24. Bentires-Alj M, Barbu V, Fillet M, et al. NF-kappaB transcription factor induces drug resistance through MDR1 expression in cancer cells. *Oncogene*. 2003;22(1):90-97. doi:10.1038/sj.onc.1206056
25. Gui Y, Qian X, Ding Y, et al. c-Fos regulated by TMPO/ERK axis promotes 5-FU resistance via inducing NANOG transcription in colon cancer. *Cell Death Dis*. 2024;15(1):61. doi:10.1038/s41419-024-06451-w
26. Jin K, Sukumar S. HOX genes: major actors in resistance to selective endocrine response modifiers. *Biochim Biophys Acta*. 2016;1865(2):105-110. doi:10.1016/j.bbcan.2016.01.003
27. Ren P, Wang JY, Zeng ZR, et al. A novel hypoxia-driven gene signature that can predict the prognosis and drug resistance of gliomas. *Front Genet*. 2022;13:976356. doi:10.3389/fgene.2022.976356
28. Liu Q, Hu Z, Jiang R, Zhou M. DeepCDR: a hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics*. 2020;36(suppl 2):i911-i918. doi:10.1093/bioinformatics/btaa822
29. Thom D, Chang RS, Lannin NA, et al. Personalised selection of medication for newly diagnosed adult epilepsy: study protocol of a first-in-class, double-blind, randomised controlled trial. *BMJ Open*. 2025;15(4):e086607. doi:10.1136/bmjopen-2024-086607
30. Nadeau C, Bengio Y. Inference for the generalization error. *Machine Learning*. 2003;52(3):239-281. doi:10.1023/A:1024068626366
31. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63. doi:10.7326/M14-0697
32. Jones KEA, Howells R, Mallick AA, Paul SP, Dey I. NICE guideline review: epilepsies in children, young people and adults NG217. *Arch Dis Child Educ Pract Ed*. 2023;108(6):416-421. doi:10.1136/archdischild-2022-324427
33. Sun Y, Seneviratne U, Perucca P, et al. Generalized polyspike train: an EEG biomarker of drug-resistant idiopathic generalized epilepsy. *Neurology*. 2018;91(19):e1822-e1830. doi:10.1212/WNL.0000000000006472
34. Salmenpera TM, Symms MR, Rugg-Gunn FJ, et al. Evaluation of quantitative magnetic resonance imaging contrasts in MRI-negative refractory focal epilepsy. *Epilepsia*. 2007;48(2):229-237. doi:10.1111/j.1528-1167.2007.00918.x
35. Szaflarski M, Szaflarski JP, Privitera MD, Ficker DM, Horner RD. Racial/ethnic disparities in the treatment of epilepsy: what do we know? What do we need to know? *Epilepsy Behav*. 2006;9(2):243-264. doi:10.1016/j.yebeh.2006.05.011