

## RESEARCH ARTICLE

# Patients carrying pathogenic *SCN8A* variants with loss- and gain-of-function effects can be classified into five subgroups exhibiting varying developmental and epileptic components of encephalopathy

Joshua B. Hack<sup>1</sup>  | Joseph C. Watkins<sup>2</sup> | John M. Schreiber<sup>3</sup>  |  
Michael F. Hammer<sup>1,4</sup> 

<sup>1</sup>BIO5 Institute, University of Arizona, Tucson, Arizona, USA

<sup>2</sup>Department of Mathematics, University of Arizona, Tucson, Arizona, USA

<sup>3</sup>Epilepsy Genetics Program, Children's National Medical Center, Washington, DC, USA

<sup>4</sup>Neurology Department, University of Arizona, Tucson, Arizona, USA

## Correspondence

Michael F. Hammer, BIO5 Institute, University of Arizona, Tucson, AZ 85721, USA.

Email: [mfh@email.arizona.edu](mailto:mfh@email.arizona.edu)

## Funding information

Shay Emma Hammer Research Foundation; Neurocrine Biosciences

## Abstract

**Objective:** Phenotypic heterogeneity presents challenges in providing clinical care to patients with pathogenic *SCN8A* variants, which underly a wide disease spectrum ranging from neurodevelopmental delays without seizures to a continuum of mild to severe developmental and epileptic encephalopathies (DEEs). An important unanswered question is whether there are clinically important subgroups within this wide spectrum. Using both supervised and unsupervised machine learning (ML) approaches, we previously found statistical support for two and three subgroups associated with loss- and gain- of- function variants, respectively. Here, we test the hypothesis that the unsupervised subgroups (U1–U3) are distinguished by differential contributions of developmental and epileptic components.

**Methods:** We predicted that patients in the U1 and U2 subgroups would differ in timing of developmental delay and seizure onset, with earlier and concurrent onset of both features for the U3 subgroup. Standard statistical procedures were used to test these predictions, as well as to investigate clinically relevant associations among all five subgroups.

**Results:** Two-population proportion and Kruskal–Wallis tests supported the hypothesis of a reversed order of developmental delay and seizure onset for patients in U1 and U2, and nearly synchronous developmental delay/seizure onset for the U3 (termed DEE) subgroup. Association testing identified subgroup variation in treatment response, frequency of initial seizure type, and comorbidities, as well as different median ages of developmental delay onset for all five subgroups.

**Significance:** Unsupervised ML approaches discern differential developmental and epileptic components among patients with *SCN8A*-related epilepsy. Patients in U1 (termed developmental encephalopathy) typically gain seizure control yet rarely experience improvements in development, whereas those in U2 (termed epileptic encephalopathy) have fewer if any developmental impairments despite difficulty in achieving seizure control. This understanding improves prognosis and clinical management and provides a framework

to discover mechanisms underlying variability in clinical outcome of patients with SCN8A-related disorders.

**KEYWORDS**

disease subphenotype, ion channel function, machine learning classifier, pediatric epilepsy

## 1 | INTRODUCTION

Phenotypic heterogeneity is recognized as a major barrier in efforts to improve clinical care and outcomes of patients with a range of diseases, including pediatric epilepsies. Hundreds of pathogenic variants in the *SCN8A* gene, which encodes the Na<sub>v</sub>1.6 voltage-gated sodium channel, underlie a wide disease spectrum that ranges from neurodevelopmental delays without seizures to severe developmental and epileptic encephalopathy (DEE).<sup>1,2</sup> One axis of variation is explained by different functional effects of loss-of-function (LOF) versus gain-of-function (GOF) variants on the Na<sub>v</sub>1.6 channel. This still leaves a broad range of clinical presentations for individuals with GOF variants and those of mixed effect.<sup>1-4</sup> An important unanswered question is whether there are clinically important patient subgroups within this wide phenotypic landscape. Individuals carrying different recurrent GOF variants are known to vary in disease course, and possibly in response to antiseizure medications (ASMs).<sup>1,5</sup> Historically, patients with GOF variants have been viewed along a mild-moderate-severe gradient in which the extent of intellectual disability (ID) and earlier age at seizure onset are markers of more severe outcome.<sup>2,6-9</sup> However, this classification may not effectively subdivide patients into clinically meaningful subgroups that inform diagnosis, prognosis, and medical care.

To test latent phenotypic structure among patients with GOF variants in the International SCN8A Registry database, Hack et al.<sup>10</sup> utilized two machine learning (ML) algorithms. Both the supervised and unsupervised approaches yielded statistical support for three distinct subgroups; however, the distinguishing features in each approach were not concordant. The supervised approach categorized individuals into three clusters (S1, S2, and S3) of increasing membership size ( $n = 17, 44,$  and  $119,$  respectively). Features such as age at seizure onset, developmental quotient (DQ), and seizure freedom were correlated, decreasing in an ordered manner from S1 to S3, with the highest level of seizure freedom in S1 (35.3%). These clusters supported the historical view of mild (S1), moderate (S2), and severe (S3) DEE subgroups.<sup>1,9,11</sup> The unsupervised approach yielded three

**Key points**

- Large-scale phenotypic variability among individuals with pathogenic *SCN8A* variants presents challenges for clinicians in identifying disease subphenotypes.
- Machine learning methods applied to clinical data in the International SCN8A Registry identified five subgroups within the *SCN8A* disease spectrum that vary in contribution of developmental and epileptic impairment.
- Subgroup variation in treatment responses, initial seizure types, and comorbidities has implications for improved prognosis and clinical management.

clusters (U1, U2, and U3, with membership size  $n = 24, 43, 113,$  respectively), in which DQ was highest in U2, whereas 100% of members in U1 experienced seizure freedom.

This discordance led us to hypothesize that the three clusters identified in the unsupervised approach represent a phenotypic landscape in which U1 is primarily a developmental encephalopathy (DE), U2 is primarily an epileptic encephalopathy (EE), and U3 represents the customary severe DEE. The rationale for this terminology is based on the finding that (1) U1 has a higher average age at seizure onset and a high rate of seizure freedom, yet has a lower average DQ; (2) U2 has an intermediate age at seizure onset and a low rate of seizure freedom, yet has a lower rate of developmental delay; and (3) U3 has an early age at seizure onset, low average DQ, and no reported cases of seizure freedom. Given this hypothesis, we predicted that individuals in the U1 subgroup would experience developmental delays prior to seizure onset, whereas those in the U2 subgroup would experience seizure onset prior to developmental delay. Those in the U3 subgroup would be expected to have nearly simultaneous developmental delay and seizure onset. Here, we test these predictions and investigate the extent to which clinically relevant associations differ among subgroups, including seizure types and response to ASMs.

## 2 | MATERIALS AND METHODS

### 2.1 | Registry cohort

Subjects in this study are participants in the International SCN8A Patient Registry (Registry) research study.<sup>12</sup> To be included, caregivers were required to provide an official genetic test report demonstrating the participant had an *SCN8A* variant classified as pathogenic or likely pathogenic. Inclusion also required participants to have provided sufficient data regarding a variety of clinical features such as those shown in Table 1. This study was active from 2014 to present, and the data were accessed for research purposes in June 2023.

### 2.2 | Ethics

The study was approved by the University of Arizona Institutional Review Board (#1603487278), and all

caregivers of individuals with SCN8A-related disorders (SCN8A-RDs) consented to participate prior to filling out the Registry questionnaire.

### 2.3 | Assignment of variant function

Functional effects reported in published electrophysiological studies performed in heterologous expression systems were used to classify variants as having GOF or LOF properties, and/or variants were classified as either GOF/LOF by utilizing the random forest classifier described by Hack et al.<sup>10</sup> Briefly, variants with a probability of LOF ( $\text{prob}[\text{LOF}] \leq .3$ ) were considered GOF, whereas those with a ( $\text{prob}[\text{LOF}] > .7$ ) were considered LOF variants. Variants with ( $\text{prob}[\text{LOF}]$ ) between .3 and .7 were considered undetermined and excluded from the study. Filtering for inclusion criteria and variant classification resulted in 180 and 50 individuals carrying variants with GOF or LOF properties, respectively.

**TABLE 1** Summary of clinical features associated with SCN8A-related disorder subgroups.

Subgroup	LOF		GOF		
	LOF–	LOF+	DE	EE	DEE
Age at seizure onset (see text for median ages)	NA	>10 months	5–10 months	4–7 months	<4 months
First seizure type(s)	NA	ABS	BTC, FOC, IS	BTC	BTC, FOC, TON, IS; >2 seizure types
Prolonged seizure freedom	NA	Possible	Possible	Difficult	Unlikely
Effects of seizure freedom on development	NA	Unknown	Continues	Improves	NA
Age at DD onset	12–48 months	>9 months	3–5 months	>8 months	<5 months
Developmental delay	Mild	Mild–moderate	Moderate–severe	Neurotypical–mild	Moderate–severe
First clinical symptom	DD	DD	DD	Epilepsy	Epilepsy and DD
Common comorbidities	Behavioral disorders	Ataxia, hypotonia, autistic features, behavioral disorders, movement disorders	Ataxia, hypotonia, autistic features	Ataxia, hypotonia, autistic features	Nonambulatory, lack of speech, G-tube fed, ataxia, hypotonia, scoliosis, pneumonia
Differentially reported ASMs	NA	ETX, LEV, CLB	SCBs and others	CBZ, OXC	CBZ, OXC, PHT, LCM
Example variant	A1622D, P1719R	G1451S, truncating	N1877S, G1475R	R1617Q, A874T	R850Q, R1872W

*Note:* Entries represent approximate values and most prevalent characteristics rather than comprising an exhaustive list. Features/ASMs that did not show differentiation among subgroups may not be listed.

Abbreviations: ABS, absence; ASM, antiseizure medication; BTC, bilateral tonic–clonic; CBZ, carbamazepine; CLB, clobazam; DD, developmental delay; DE, developmental encephalopathy; DEE, developmental and epileptic encephalopathy; EE, epileptic encephalopathy; ETX, ethosuximide; FOC, focal; GOF, gain of function; IS, infantile spasms; LCM, lacosamide; LEV, levetiracetam; LOF, loss of function; NA, not available; OXC, oxcarbazepine; PHT, phenytoin; SCB, sodium channel blocker; TON, tonic.

## 2.4 | Onset comparisons

The age at seizure onset and developmental delay onset were obtained for each individual. seizure onset–developmental delay onset was calculated for all individuals, a metric that reflected the difference in timing of seizure *versus* developmental delay onset. Two-population proportion tests were conducted comparing frequency of seizure–developmental delay, and Kruskal–Wallis tests followed by Dunn tests were conducted on the average gap between these two onset ages for U1, U2, and DEE.

## 2.5 | Medication comparisons

Data on 32 prescribed ASMs were analyzed, including a caregiver ranking for best and worst ASM. Also analyzed were data on the reasons ASMs were deemed beneficial and detrimental, including effects on seizure frequency, seizure duration, alertness, and adverse side effects. After constructing contingency tables with the number of individuals reported with a particular response to an ASM *versus* the remaining number of individuals in a given subgroup, a series of  $4 \times 2$  (i.e., all subgroups with seizures) and  $3 \times 2$  (i.e., GOF-associated subgroups only) exact tests were performed to test for statistical significance. Similar analyses were performed for initial seizure types and comorbidities. Patients with SCN8A-RDs infrequently undergo testing for ID, and therefore may not be diagnosed with ID, or may be considered to have global developmental delay and are not yet diagnosed with ID. Therefore, ID features as reported by caregivers in the Registry are likely to be underreported relative to clinically assessed ID. We note that patients without recorded ID had uncontrolled seizures, which in the SCN8A population are only found in patients with encephalopathy. All statistical procedures were conducted in RStudio v4.2.2.

# 3 | RESULTS

## 3.1 | Study cohort

This study included 230 individuals from the International SCN8A Registry. The median age at completing the Registry was 59 months. The median age at initial diagnosis was 5 months, whereas the median age at SCN8A-RD diagnosis was 22 months. This cohort was comprised of 180 (78.3%) individuals with GOF variants and 50 (21.7%) with LOF variants; 219 (95.2%) individuals possessed missense variants, and 11 (4.8%) possessed truncating variants. Additional demographic data are described in Andrews et al.<sup>12</sup>

## 3.2 | GOF subphenotypic categories

This study included all individuals except one with missing developmental data reported in Hack et al.<sup>10</sup> ( $n = 179$ ): 23 individuals in U1, 43 in U2, and 113 in U3. Of the three categories, U2 was the most variable in age at developmental delay and seizure onset. The median age at seizure onset was  $5.04 \pm 2.8$  months,  $4.72 \pm 3.51$  months, and  $2.88 \pm 2.27$  months for U1, U2, and U3, respectively. A Kruskal–Wallis test resulted in a  $p$ -value of  $2.12 \times 10^{-4}$ , and Dunn tests identified DEE as distinct from U1 ( $p = 2.04 \times 10^{-3}$ ) and U2 ( $p = 5.47 \times 10^{-3}$ ). The average age at developmental delay onset was  $4.00 \pm 6.78$  months,  $32.72 \pm 19.03$  months, and  $3.16 \pm 1.77$  months for U1, U2 (when developmental delay occurred), and U3, respectively. Kruskal–Wallis testing resulted in a  $p$ -value of  $1.32 \times 10^{-21}$ , and Dunn tests resulted in significant differences between U2 and U1 ( $p = 6.44 \times 10^{-13}$ ) and U3 ( $p = 3.02 \times 10^{-20}$ ). There were 35 individuals reporting no significant developmental delays: none in U1, 33 in U2, and two in U3 (two-sample proportion test U1:U2  $p = 7.94 \times 10^{-9}$ ). However, these individuals likely are encephalopathic, because encephalopathy is essentially always seen in people with intractable seizures due to SCN8A mutations.<sup>13</sup>

Fifty individuals possessing LOF variants were included in the dataset, with 31 having experienced seizures (LOF+) and the other 19 having never experienced seizures (LOF–). The median age at seizure onset for LOF+ was 18 months. The average age at developmental delay onset was  $7.1 \pm 5.7$  months for LOF– and  $6.44 \pm 4.0$  months for LOF+. Of those in LOF+, 13 of 31 achieved age-appropriate development. The study population contains no individuals with self-limited (familial) infantile epilepsy (SeL[F]IE).<sup>14</sup> All individuals in the study continue to have seizures and/or have encephalopathy at the time of reporting.

## 3.3 | Developmental delay and seizure onset ages

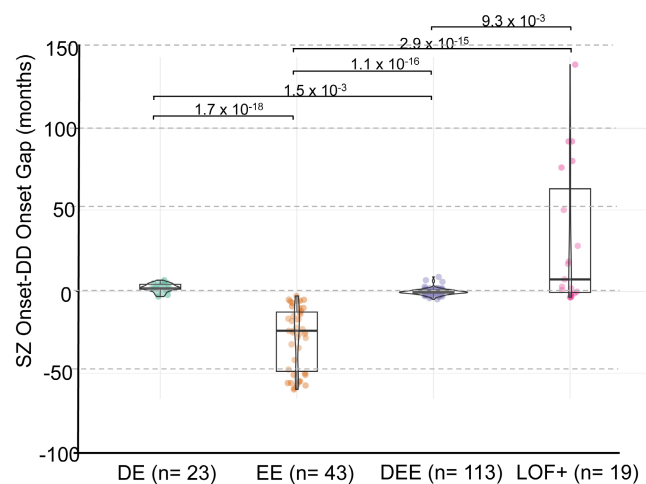
Table S1 shows the number of individuals experiencing seizure onset prior to developmental delay onset and the average seizure–developmental delay onset gap for each classification. The number with seizure onset before developmental delay onset was five of 23 for U1, 43 of 43 for U2, and 81 of 113 for U3. Two-population proportion tests conducted on the frequency of seizure onset prior to developmental delay onset to determine whether the three classifications were distinct from each other showed  $p$ -values of  $<1 \times 10^{-5}$ ,  $5 \times 10^{-5}$ , and  $<1 \times 10^{-5}$ , comparing U3 to U1, U3 to U2, and U2 to U1, respectively. Of those

in LOF+ who experienced developmental delay, 13 of 18 (72.2%) experienced developmental delay onset prior to seizure onset.

The average seizure–developmental delay onset gap for U1 was  $2.39 \pm 2.40$  months, for U2 was  $-28.0 \pm 18.66$  months, and for U3 was  $-.2 \pm 1.86$  months. In these cases, negative values indicate seizure onset prior to developmental delay onset (Table S1). A Kruskal–Wallis test resulted in a  $p$ -value of  $8.88 \times 10^{-25}$ , with Dunn tests identifying each subgroup to be significantly different from each other (Figure 1). The average seizure–developmental delay onset gap for LOF+ was  $31.2 \pm 42.2$  months.

### 3.4 | ASM responses

ASMs that were most frequently prescribed included leviracetam (LEV;  $n = 167$ ), oxcarbazepine (OXC;  $n = 118$ ), clobazam (CLB;  $n = 118$ ), topiramate ( $n = 107$ ), and phenobarbital ( $n = 98$ ). The frequency of best and worst ASMs reported for all patients is shown in Figure 2A. As expected, the sodium channel blockers (SCBs) were reported as the most beneficial ASMs for individuals in U1, U2, and U3 (Figure 2A). Of all 32 ASMs, three were found to be differentially reported as the best ASM for GOF-associated subgroups versus the LOF+ subgroup (Table S2, Figure 2B left). OXC was the most frequently reported best ASM for U1, U2, and U3 subgroups ( $4 \times 2$  exact test  $p$ -value =  $3.2 \times 10^{-2}$ ). Although rates of success with CBZ differed between LOF+ and the GOF subgroups, it was not reported as the best ASM as often as OXC and



**FIGURE 1** Violin plot with Kruskal–Wallis and Dunn tests results for seizure (seizure) onset–developmental delay onset gap statistic by subgroup. Probability values are shown above the comparison bars. DD, developmental delay; DE, developmental encephalopathy; DEE, developmental and epileptic encephalopathy; EE, epileptic encephalopathy; LOF, loss of function.

therefore did not reach statistical significance as a differentiating ASM. Phenytoin (PHT) and lacosamide were differentially reported for GOF-associated subgroups ( $4 \times 2$  exact test  $p$ -value =  $2.0 \times 10^{-3}$  and  $1.2 \times 10^{-2}$ , respectively) and were more frequently reported as the best ASMs for U3 ( $3 \times 2$  exact test  $p$ -value =  $1.4 \times 10^{-2}$  and  $1.1 \times 10^{-2}$ , respectively). The latter was most apparent in the U3 versus U2 comparison ( $2 \times 2$  exact test  $p$ -value =  $7.8 \times 10^{-3}$  and  $5.0 \times 10^{-3}$ , respectively). OXC was reported to have an increased rate of seizure frequency reduction in the GOF-associated subgroups compared with LOF+ ( $4 \times 2$  exact test  $p$ -value =  $8.0 \times 10^{-3}$ ), and PHT was reported to both reduce seizure frequency and duration ( $4 \times 2$  exact test  $p$ -value =  $3.0 \times 10^{-3}$  and  $1.0 \times 10^{-3}$ , respectively; Table S2, Figure S1). PHT also had a higher reported rate of reducing frequency and duration in U3 versus U1 and U2 subgroups ( $3 \times 2$  exact test  $p$ -value =  $1.6 \times 10^{-3}$  and  $8.0 \times 10^{-3}$ , respectively; Table 1).

Comparisons among subgroups revealed that ethosuximide (ETX) and LEV were more frequently reported as the best ASMs for LOF+ ( $4 \times 2$  exact test  $p$ -value =  $1.5 \times 10^{-2}$  and  $6.0 \times 10^{-3}$ , respectively), whereas valproic acid (VPA) was more frequently reported as the worst ASM for LOF+ (Table S2, Figure 2B, right). Interestingly, LEV was frequently reported as the worst ASM for all four subgroups. ETX and LEV both reduced seizure frequency in LOF+ ( $4 \times 2$  exact test  $p$ -value =  $2 \times 10^{-3}$  and  $7 \times 10^{-3}$ , respectively; Table S2, Figure S1). CLB and ETX were reported to increase alertness in LOF+ ( $4 \times 2$  exact test  $p$ -value =  $1.8 \times 10^{-2}$  and  $4.5 \times 10^{-2}$ , respectively). Lamotrigine was reported to decrease alertness in U2 ( $3 \times 2$  exact test  $p$ -value =  $4.4 \times 10^{-2}$ ) and increase side effects in LOF+ ( $4 \times 2$  exact test  $p$ -value =  $1.0 \times 10^{-2}$ ), and VPA was reported to decrease alertness and increase side effects in LOF+ ( $4 \times 2$  exact test  $p$ -value =  $4.3 \times 10^{-2}$  and  $4.0 \times 10^{-3}$ , respectively; Table S2).

### 3.5 | Seizure patterns and types

In the case of U1, seizure onset is generally between 5 and 10 months (Table S3). Typically, a single seizure type presents initially, whether bilateral tonic–clonic (BTC), infantile spasms (IS), or focal (FOC; Figure 2C). Individuals in U2 experience initial seizure onset between 4 and 7 months (Table 1). Initial seizures are ordinarily limited to only one or two types, with BTC predominating (Figure 2C). seizure onset for individuals in U3 is prior to 4 months, often with multiple seizure types presenting early. Similar to other subgroups, BTC is the most common; however, FOC, tonic (TON), and IS are also present relatively frequently (Table S3, Figure 2C).

Seizure types were not reported at equal frequency across GOF-associated subgroups (Table S3, Figure 2C).



BTC seizures were reported approximately twice as frequently in the U2 and U3 subgroups ( $3 \times 2$  exact test  $p$ -value =  $1.9 \times 10^{-2}$ ), whereas FOC seizures were more frequently reported in U3 by a factor of more than two ( $3 \times 2$  exact test  $p$ -value =  $2.0 \times 10^{-3}$ ). IS are reported at a rate more than five times higher in U1 and U3 compared with U2 ( $3 \times 2$  exact test  $p$ -value =  $1.0 \times 10^{-3}$ ). In

addition, TON, myoclonic, and absence (ABS) seizures are reported to occur more frequently in the U3 subgroup ( $3 \times 2$  exact test  $p$ -value =  $1.0 \times 10^{-3}$ ,  $5.0 \times 10^{-2}$ ,  $3.0 \times 10^{-2}$ , respectively). For LOF+, initial seizure onset is typically after 10 months, with generalized ABS or atypical ABS seizures dominating as the most common initial seizure type at a rate that is statistically significantly higher than

**FIGURE 2** Relative efficacy, beneficial, and detrimental effects of medication by subgroup, and comparison of seizure types and comorbidities differing among SCN8A-related disorder (SCN8A-RD) subgroups. (A) Antiseizure medications (ASMs) reported as best or worst for (left) individuals with gain-of-function (GOF) variants and (right) individuals with loss-of-function (LOF) variants. (B) ASMs reported as best (left) and worst (right) that differed significantly among subgroups. (C) Comparison of seizure types (left) and comorbidities (right) differing among SCN8A-RD subgroups. Values represent percentage of reports normalized by subgroup sample sizes. On the left, percentage values represent the number of reports divided by subgroup sample sizes. Significant exact test *p*-values are indicated by (1) letter G above bar for GOF *versus* LOF, (2) letter L above bar for LOF *versus* GOF, and (3) asterisk for ASM among GOF subgroups. Significant exact test *p*-values are indicated by letter L above bar for LOF *versus* GOF and asterisk for seizure type that differs among GOF subgroups. On the right, prevalence of common comorbidities (>15%) that differ among SCN8A-RD subgroups is shown. Intellectual disability (ID), which may be underreported in the study cohort (see Materials and Methods), is significantly more prevalent in developmental and epileptic encephalopathy (DEE) compared with other subgroups. It is also significantly less common in LOF– compared with LOF+. Hypotonia (HYP) is relatively common in all subgroups. Autistic features (AUT) are more common in LOF+ compared with all subgroups, including LOF–. LOF– exhibits significantly lower levels of ID, ataxia (ATX), pain tolerance (PT), unsafe swallow (US), feeding issues (e.g., G-tube fed, GT), increased sweating (SW), and scoliosis (SC). US, SW, and SC are significantly elevated in DEE compared with the developmental encephalopathy (DE) and epileptic encephalopathy (EE) subgroups, and GT is elevated in EE and DEE. Symbols: ~, approximately equivalent among subgroups; asterisk up arrow, significantly elevated among GOF subgroups; L up arrow, significantly elevated compared with all other subgroups; L down arrow, significantly less frequent compared with all other subgroups. ABS, absence; ACTH, adrenocorticotropic hormone; ATO, atonic; BTC, bilateral tonic-clonic; CBD, cannabidiol; CBZ, carbamazepine; CCB, calcium channel blocker; CLB, clobazam; CLZ, clonazepam; EPI, Epidiolex; ESL, eslicarbazepine; ETX, ethosuximide; FOC, focal; GABA,  $\gamma$ -aminobutyric acidergic; GBP, gabapentin; IS, infantile spasms; LCM, lacosamide; LEV, levetiracetam; LTG, lamotrigine; MULT, multiple mechanisms of action; MYO, myoclonic; OXC, oxcarbazepine; PBT, phenobarbital; PHT, phenytoin; PRD, prednisolone; RFM, rufinamide; SCB, sodium channel blocker; STP, stiripentol; STR, steroid; SVP2, synaptic vesicle protein 2A; TON, tonic; TPM, topiramate; VGB, vigabatrin; VPA, valproate; ZNS, zonisamide.

for the GOF-associated subgroup ( $4 \times 2$  exact test *p*-value =  $< 1.0 \times 10^{-5}$ ; Table 1).

### 3.6 | Comorbidities

The U1 subgroup typically presents with ID, ataxia, autistic features, and hypotonia. Individuals in U2 typically experience developmental delay after seizure onset, if at all (Table S4, Figure 2C). ID and hypotonia are relatively common, with autistic features, ataxia, and pain tolerance reported less frequently. Although relatively infrequent, feeding issues (e.g., G-tube fed) are more commonly reported in U2 than in U1, and most frequently reported in the case of U3. ID is statistically significantly more frequent in U3 compared with U1 and U2, as is unsafe swallow and scoliosis. Autistic features, ataxia, and pain tolerance are also relatively frequently reported in U3. In the case of LOF+, autistic features are more frequently reported at a statistically significant level compared with all other subgroups. Also frequently reported are ID, hypotonia, ataxia, and pain tolerance. Table 1 provides a summary of the abovementioned results.

## 4 | DISCUSSION

Given the phenotypic heterogeneity associated with SCN8A-RDs and the numerous pathogenic variants in the gene with both LOF and GOF effects, clinicians have been

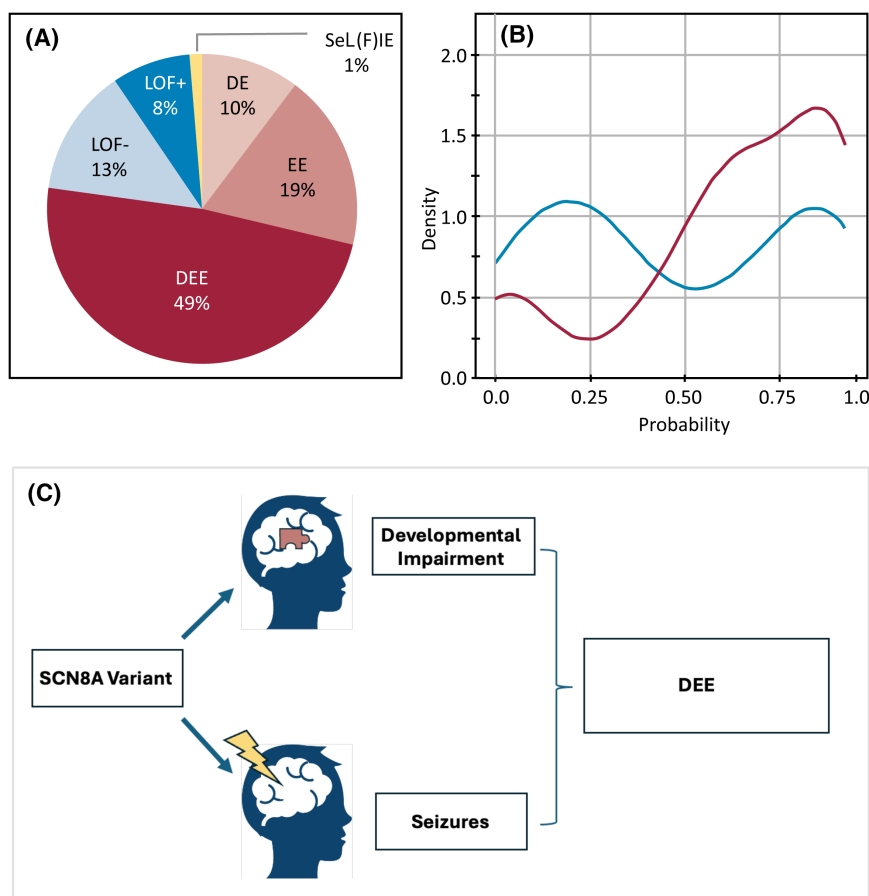
challenged to provide accurate diagnosis and prognosis and a course of effective treatment.<sup>1,9</sup> It is clear that the LOF and GOF subgroups differ in biological terms, given the different effects these variant types have on channel function.<sup>2,3,15</sup> This distinction carries important implications for treatment, as patients with GOF variants often benefit from SCBs, whereas SCBs tend to aggravate symptoms of patients with LOF variants.<sup>10,16</sup> Patients with GOF variants comprise a large portion of the phenotypic spectrum that has been historically viewed along a mild-moderate-severe gradient,<sup>2,7-9</sup> and our supervised ML models have generally supported this phenotypic structure.<sup>17</sup> However, the question of whether these patients can be classified into clinically meaningful subgroups has not been adequately addressed. Here, we tested the hypothesis proposed in Hack et al.<sup>17</sup> that three clinically distinct subgroups initially referred to as U1, U2, and U3 exist within the GOF population. To this end, we *a priori* predicted that individuals in the U1 subgroup experienced developmental delay onset prior to seizure onset, those in the U2 subgroup experienced seizure onset prior to developmental delay onset, and those in the U3 subgroup experienced both seizure and developmental delay onset nearly concurrently. The results of statistical procedures described above support these predictions (Figure 1, Table S2) and demonstrate that there are three high-level “partially ordered” subgroups, meaning that these subgroups neither conform to a typical mild to severe spectrum nor do they exist as separate, unordered categories. Altogether, the results support the existence of at least

five clinically relevant subgroups within the SCN8A-RD phenotypic spectrum: two associated with LOF variants and three with GOF variants (Table 1). Given these phenotypic patterns, we refer to the LOF subgroups as LOF+ and LOF−, as they respectively present with or without seizures, and the U1, U2, and U3 GOF subgroups as DE, EE, and DEE, as they respectively present primarily with DE, EE, or a combination of both (DEE) components contributing to encephalopathy (Figure 3A). It is important to note that it has been difficult to distinguish mildly and moderately affected patients in clinical settings,<sup>1,9,11</sup> and that our own ML analyses provide a cleaner separation of DE and EE subgroups compared with mild and moderate

subgroups. This can be seen in the more distinct bimodal pattern separating DE from EE, as displayed in Figure 3B.

#### 4.1 | DEE concept

The concept of EE was modified in 2017 by the addition of the term “developmental” (i.e., DEE) to acknowledge the growing number of genetic epilepsies in which patients display both delayed development and very active epileptiform abnormalities.<sup>19–21</sup> In 2022, the DEE concept was further clarified in recognition of the finding that developmental delay precedes seizure onset in some conditions,



**FIGURE 3** (A) Frequency of the five subgroups included in this study, including an estimated frequency of self-limited (familial) infantile epilepsy (SeL[F]IE) based on participant composition in the International SCN8A Registry (see text). (B) Ability to distinguish among classifications for the mild/moderate *versus* the developmental encephalopathy (DE)/epileptic encephalopathy (EE) subgroups using supervised and unsupervised learning. For every individual, in the mild/moderate category, we give the probability that the individual is classified as moderate. Thus, the individuals classified as mild should have a low probability of classification as moderate (red peak on the left). Correspondingly, individuals classified as moderate should have a high probability of classification as moderate (red peak on the right). Similarly, for every individual, in the DE/EE category, we give the probability that the individual is classified as EE. Thus, the individuals classified as DE should have a low probability of the EE classification (blue peak on the left). Correspondingly, individuals classified as EE should have a high probability of the EE classification (blue peak on the right). Note that the bimodal distribution is more distinct in the unsupervised classification scheme, showing that this approach is more effective in distinguishing groups than the approach based on supervised learning. (C) Conceptualization of developmental and epileptic encephalopathies (DEEs) as a partially ordered model with the DEE subgroup having both DE and EE components, and the DE and EE subgroups as distinct but not separate unordered categories (see Lin et al.<sup>18</sup>). LOF, loss of function.

whereas epileptic activity leads to cognitive and behavioral deterioration in other conditions.<sup>18,22</sup> In this model, epilepsy and developmental impairment are viewed as epiphenomena of the underlying pathology (i.e., implying that developmental impairments are not necessarily caused by the seizures or interictal epileptic activity per se<sup>18,21</sup>; Figure 3C). Interestingly, a recent rapid genome sequencing cohort of diverse infantile onset epilepsy disorders found that 25% of patients had developmental delay before seizure onset and 31% had developmental plateau or regression following seizure onset.<sup>23</sup> Consistent with this view, we find that 36.9% of our SCN8A-RD cohort with GOF variants experience developmental delay before seizure onset or vice versa. In the remaining patients in our cohort, there appeared to be concurrent onset of seizures and developmental delays, 29% of whom experienced both seizure and developmental delay onset in the first 3 months of life. Similarly, phenotypic patterns in patients with LOF variants support the hypothesis of a DE independent of seizure activity, given the extent of developmental impairment in both subgroups and the later seizure onset in the case of LOF+.

## 4.2 | Pathophysiological implications

These results support the aforementioned hypothesis that SCN8A-related epilepsy and developmental impairment represent secondary outcomes (epiphenomena) of a common underlying pathology (Figure 3C).<sup>18</sup> Our findings of the significant differences among the DE, EE, and DEE subgroups in seizure patterns, developmental impairment, comorbidities, and responses to ASMs suggest different clinical trajectories.<sup>24,25</sup> In the case of DE, seizure freedom is not associated with developmental improvement, *contra* findings with EE. This suggests an underlying developmental pathophysiology not necessarily related to seizure burden, whereas developmental impairments in EE may be directly related to the impact of seizures themselves.<sup>21</sup> In the case of DEE, developmental impairments may be exacerbated by seizures but not remit when seizures are partially controlled, suggesting a more complex underlying pathophysiology. Interestingly, developmental impairment is also a main feature in the LOF subgroups, which points to the possibility that depolarization block and action potential failure may underlie developmental impairments in patients carrying variants with strong GOF properties.<sup>3,26</sup> This highlights the need to better understand the mechanisms underlying epilepsy and developmental impairments even in cases of a single gene etiology.<sup>27</sup>

There are several non-mutually exclusive factors that may contribute to different clinical trajectories among subgroups. There are associations between particular

GOF variants and phenotypic subgroups.<sup>1,2,5,17</sup> As expected, there are a host of variants—both truncating and missense—exclusively associated with LOF subgroups. A recurrent missense variant is primarily found in individuals with LOF+ (G1451S; Table S3). Other non-mutually exclusive biological factors that could explain different disease trajectories among subgroups include modifier genes,<sup>28,29</sup> genetic background effects,<sup>30</sup> and cellular signaling and pathway variation.<sup>31</sup>

## 4.3 | Comparisons with the clinical consensus

Recently, a modified Delphi method was utilized to produce a clinical consensus on current practices for diagnosis and treatment<sup>32</sup> and prognosis and comorbidities<sup>33</sup> associated with the full range of SCN8A-RD phenotypes. Relatively high consensus (81%–97%) was achieved for a phenotypic landscape comprised of severe DEE, mild/moderate DEE, SeL(F)IE, neurodevelopmental disability (NDD) with generalized epilepsy, and NDD without epilepsy. Note that mild and moderate patients were lumped into a single subgroup, and SeL(F)IE comprised a third GOF subgroup. The unsupervised classification schema concurs with the consensus schema in identifying DEE as the most prevalent subgroup (labeled severe DEE in the consensus), as well as in classifying patients with LOF variants into two subgroups (here labeled LOF+ and LOF–; Table S5). The two schemas also agree in the co-occurrence of seizure and developmental delay onset in the DEE subgroup, and that developmental delay is the primary presenting feature of in the LOF categories. There is also general agreement in a pattern of increasing age at seizure onset as severity of outcome increases (Table S5). However, age at seizure onset is not predictive in itself for important outcome measures such as developmental quotient, comorbidities, seizure types, and expected degree of seizure freedom.

As discussed above, a major advantage of the unsupervised approach is that it differentiates patients who may otherwise be classified as mild/moderate into two distinct subgroups, and as such represents an advancement in the classification of patients carrying variants with GOF properties. Importantly, patients in the developmental delay subgroup typically continue to be challenged with developmental impairments despite some level of seizure control, whereas those in the EE subgroup may struggle with seizure control and yet achieve developmental milestones in a manner more similar to neurotypical individuals. These results also reveal that the less severe SCN8A phenotypes may have variable contributions from the “developmental” and “epileptic” components of their pathophysiology

that can influence treatment response and developmental outcome. The timing of seizure onset and developmental delay onset may help to predict these outcomes.

#### 4.4 | Significance and limitations

A major shortcoming in the classification schema used in the consensus process, as well as that supported in the supervised approach, is that they are retrospective in nature, relying on the outcomes of disease progression to classify patients into phenotypic subgroups. In the current modeling approach,<sup>17</sup> we identify and rank the importance of features that may not have been previously recognized as predictive of outcome and that can form the basis for distinct subgroups. The use of unsupervised ML techniques in classification procedures also helps to minimize confirmation bias and aids in classifying patients into more clinically meaningful groups (i.e., clusters of patients with shared combinations of traits such as seizure freedom, developmental quotient, response to ASMs, and variant function<sup>34</sup>; Table 1). It is important to note that these subgroups are partially ordered and do not exist as separate, unordered categories (Figure 3C). This advances our understanding of SCN8A-RD from a retrospective to a prospective classification. The next step is the validation of this schema using distinct datasets and eventual evaluation on the basis of a prospective natural history study. In sum, this work supports a reappraisal of single gene disorders classified broadly as DEEs, where distinct EE, DE, and DEE subgroups may exist, yielding important insights into treatment, genetic counseling, and prognosis.<sup>18</sup>

#### AUTHOR CONTRIBUTIONS

**Joshua B. Hack:** Study design; statistical analysis and interpretation of data; drafting the manuscript. **Joseph C. Watkins:** Statistical analysis and interpretation of data; drafting the manuscript. **John M. Schreiber:** Interpretation of data; drafting the manuscript. **Michael F. Hammer:** Study design; analysis and interpretation of data; drafting the manuscript.

#### ACKNOWLEDGMENTS

We thank Anne Berg for comments and Neurocrine Biosciences and the Shay Emma Hammer Research Foundation for funding. We are particularly grateful to all the families that participated in the International SCN8A Registry research study.

#### FUNDING INFORMATION

This study was funded by Neurocrine Biosciences and the Shay Emma Hammer Research Foundation.

#### CONFLICT OF INTEREST STATEMENT

The authors report no disclosures relevant to the article. 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.

#### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Supplementary Information; additional data are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

#### ORCID

Joshua B. Hack  <https://orcid.org/0000-0003-0469-4559>

John M. Schreiber  <https://orcid.org/0000-0003-0615-2497>

Michael F. Hammer  <https://orcid.org/0000-0003-0172-429X>

#### REFERENCES

1. Hammer MF, Xia M, Schreiber JM. In: Adam MP, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al., editors. SCN8A-related epilepsy and/or neurodevelopmental disorders. Seattle (WA): GeneReviews(R); 2023.
2. Johannesen KM, Liu Y, Koko M, Gjerulfson CE, Sonnenberg L, Schubert J, et al. Genotype-phenotype correlations in SCN8A-related disorders reveal prognostic and therapeutic implications. *Brain*. 2022;14(145):2991–3009.
3. Liu Y, Schubert J, Sonnenberg L, Helbig KL, Hoei-Hansen CE, Koko M, et al. Neuronal mechanisms of mutations in SCN8A causing epilepsy or intellectual disability. *Brain*. 2019;1(142):376–90.
4. Lyu H, Bosselmann CM, Johannesen KM, Koko M, Ortigoza-Escobar JD, Aguilera-Albesa S, et al. Clinical and electrophysiological features of SCN8A variants causing episodic or chronic ataxia. *EBioMedicine*. 2023;98:104855.
5. Chung KM, Hack J, Andrews J, Galindo-Kelly M, Schreiber J, Watkins J, et al. Clinical severity is correlated with age at seizure onset and biophysical properties of recurrent gain of function variants associated with SCN8A-related epilepsy. *Epilepsia*. 2023.
6. Encinas AC, Moore IKM, Watkins JC, Hammer MF. Influence of age at seizure onset on the acquisition of neurodevelopmental skills in an SCN8A cohort. *Epilepsia*. 2019;60:1711–20.
7. Johannesen KM, Gardella E, Encinas AC, Lehesjoki AE, Linnankivi T, Petersen MB, et al. The spectrum of intermediate SCN8A-related epilepsy. *Epilepsia*. 2019;60:830–44.
8. Larsen J, Carvill GL, Gardella E, Kluger G, Schmiedel G, Barisic N, et al. The phenotypic spectrum of SCN8A encephalopathy. *Neurology*. 2015;3(84):480–9.
9. Talwar D, Hammer MF. SCN8A epilepsy, developmental encephalopathy, and related disorders. *Pediatr Neurol*. 2021;122:76–83.
10. Hack JB, Horning K, Juroske Short DM, Schreiber JM, Watkins JC, Hammer MF. Distinguishing loss-of-function and gain-of-function SCN8A variants using a random Forest classification model trained on clinical features. *Neurol Genet*. 2023;9:1–10.

11. Johannesen KM, Gardella E, Ahring PK, Moller RS. De novo SCN3A missense variant associated with self-limiting generalized epilepsy with fever sensitivity. *Eur J Med Genet.* 2022;65:104577.
12. Andrews J, Kelly Galindo M, Hack JB, Watkins JC, Conecker GA, Hammer MF, et al. The international SCN8A patient registry: a scientific resource to advance the understanding and treatment of a rare pediatric neurodevelopmental syndrome J registry. *Management.* 2023;50:4–10.
13. Khan S, Al BR. Epileptic encephalopathies: an overview. *Epilepsy Res Treat.* 2012;2012:403592.
14. van Roest A, Van de Vel A, Lederer D, Ceulemans B. The clinical and genetic spectrum in infants with (an) unprovoked cluster(s) of focal seizures. *Eur J Paediatr Neurol.* 2020;24:148–53.
15. Wagnon JL, Bunton-Stasyshyn RK, Meisler MH. Mutations of Sodium Channel SCN8A (Nav1.6) in neurological disease. In: Pitt GS, editor. *Ion channels in health and disease.* San Diego: Academic Press; 2016. p. 239–64.
16. Meisler MH, Hill SF, Yu W. Sodium channelopathies in neurodevelopmental disorders. *Nat Rev Neurosci.* 2021;22:152–66.
17. Hack JB, Watkins JC, Hammer MF. Machine learning models reveal distinct disease subgroups and improve diagnostic and prognostic accuracy for individuals with pathogenic SCN8A gain-of-function variants *Biol Open.* 2024;13:bio060286.
18. Lin JJ, Meletti S, Vaudano AE, Lin KL. Developmental and epileptic encephalopathies: is prognosis related to different epileptic network dysfunctions? *Epilepsy Behav.* 2022;131:107654.
19. Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, van Emde BW, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE commission on classification and terminology, 2005-2009. *Epilepsia.* 2022;51:676–85.
20. Dulac O. Epileptic encephalopathy. *Epilepsia.* 2001;42(Suppl 3):23–6.
21. Scheffer IE, Liao J. Deciphering the concepts behind “epileptic encephalopathy” and “developmental and epileptic encephalopathy”. *Eur J Paediatr Neurol.* 2020;24:11–4.
22. Zuberi SM, Wirrell E, Yozawitz E, Wilmschurst JM, Specchio N, Riney K, et al. ILAE classification and definition of epilepsy syndromes with onset in neonates and infants: position statement by the ILAE task force on nosology and definitions. *Epilepsia.* 2022;63:1349–97.
23. D’Gama AM, Mulhern S, Sheidley BR, Boodhoo F, Buts S, Chandler NJ, et al. Evaluation of the feasibility, diagnostic yield, and clinical utility of rapid genome sequencing in infantile epilepsy (gene-STEPS): an international, multicentre, pilot cohort study. *Lancet Neurol.* 2023;22:812–25.
24. Akerlund CAI, Holst A, Stocchetti N, Steyerberg EW, Menon DK, Ercole A, et al. Clustering identifies endotypes of traumatic brain injury in an intensive care cohort: A CENTER-TBI study. *Crit Care.* 2022;27(26):228.
25. Ranard BL, Megjhani M, Terilli K, Doyle K, Claassen J, Pinsky MR, et al. Identification of Endotypes of hospitalized COVID-19 patients. *Front Med (Lausanne).* 2021;8:770343.
26. Wengert ER, Miralles RM, Wedgwood KCA, Wagley PK, Strohm SM, Panchal PS, et al. Somatostatin-positive interneurons contribute to seizures in SCN8A epileptic encephalopathy. *J Neurosci.* 2021;3(41):9257–73.
27. Shimizu H, Morimoto Y, Yamamoto N, Tayama T, Ozswa H, Imamura A. Overlap between epilepsy and neurodevelopmental disorders: insights from clinical and genetic studies. In: Czuczwar SJ, editor. *Epilepsy.* Brisbane, Australia: Exon Publications; 2022. p. 41–54.
28. Emond MJ, Louie T, Emerson J, Chong JX, Mathias RA, Knowles MR, et al. Exome sequencing of phenotypic extremes identifies CAV2 and TMC6 as interacting modifiers of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *PLoS Genet.* 2015;11:e1005273.
29. Mersha TB, Afanador Y, Johansson E, Proper SP, Bernstein JA, Rothenberg ME, et al. Resolving clinical phenotypes into Endotypes in allergy: molecular and omics approaches. *Clin Rev Allergy Immunol.* 2021;60:200–19.
30. Hammer MF, Ishii A, Johnstone L, Tchourbanov A, Lau B, Sprissler R, et al. Rare variants of small effect size in neuronal excitability genes influence clinical outcome in Japanese cases of SCN1A truncation-positive Dravet syndrome. *PLoS One.* 2017;12:e0180485.
31. Agache I, Akdis CA. Precision medicine and phenotypes, endotypes, genotypes, regiotypes, and theratypes of allergic diseases. *J Clin Invest.* 2019;11(129):1493–503.
32. Conecker G, Xia MY, Hecker J, Achkar C, Cukiert C, Devries S, et al. Global modified Delphi consensus on diagnosis, phenotypes, and treatment of SCN8A-related epilepsy and/or neurodevelopmental disorders. *Epilepsia.* 2024;65:2322–38.
33. Conecker G, Xia MY, Hecker J, Achkar C, Cukiert C, Devries S, et al. Global modified-Delphi consensus on comorbidities and prognosis of SCN8A-related epilepsy and/or neurodevelopmental disorders. *Epilepsia.* 2024;65:2308–21.
34. Levine AR, Calfee CS. Subphenotypes of acute respiratory distress syndrome: advancing towards precision medicine. *Tuberc Respir Dis (Seoul).* 2024;87:1–11.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Hack JB, Watkins JC, Schreiber JM, Hammer MF. Patients carrying pathogenic *SCN8A* variants with loss- and gain-of-function effects can be classified into five subgroups exhibiting varying developmental and epileptic components of encephalopathy. *Epilepsia.* 2024;00:1–11. <https://doi.org/10.1111/epi.18118>