Author/s:
Abou-Khalil, B;Afawi, Z;Allen, AS;Bautista, JF;Bellows, ST;Berkovic, SF;Bluvstein, J;Burgess, R;Cascino, G;Cossette, P;Cristofaro, S;Crompton, DE;Delanty, N;Devinsky, O;Dlugos, D;Ellis, CA;Epstein, MP;Fountain, NB;Freyer, C;Geller, EB;Glauser, T;Glynn, S;Goldberg-Stern, H;Goldstein, DB;Gravel, M;Haas, K;Haut, S;Heinzen, EL;Kirsch, HE;Kivity, S;Knowlton, R;Korczyn, AD;Kossoff, E;Kuzniecky, R;Loeb, R;Lowenstein, DH;Marson, AG;McCormack, M;McKenna, K;Mefford, HC;Motika, P;Mullen, SA;J. O'Brien, T;Ottman, R;Paolicchi, J;Parent, JM;Paterson, S;Petrou, S;Petrovski, S;Owen Pickrell, W;Poduri, A;Rees, MI;Sadleir, LG;Scheffer, IE;Shih, J;Singh, R;Sirven, J;Smith, M;Smith, PEM;Thio, LL;Thomas, RH;Venkat, A;Vining, E;Von Allmen, G;Weisenberg, J;Widdess-Walsh, P;Winawer, MR

Title:
Quantitative analysis of phenotypic elements augments traditional electroclinical classification of common familial epilepsies

Date:
2019-11-01

Citation:
Abou-Khalil, B., Afawi, Z., Allen, A. S., Bautista, J. F., Bellows, S. T., Berkovic, S. F., Bluvstein, J., Burgess, R., Cascino, G., Cossette, P., Cristofaro, S., Crompton, D. E., Delanty, N., Devinsky, O., Dlugos, D., Ellis, C. A., Epstein, M. P., Fountain, N. B., Freyer, C. ,... Winawer, M. R. (2019). Quantitative analysis of phenotypic elements augments traditional electroclinical classification of common familial epilepsies. Epilepsia, 60 (11), pp.2194-2203. https://doi.org/10.1111/epi.16354.

Article type : Full length original research paper

Accept Date: 04-Sep-2019

Quantitative analysis of phenotypic elements augments traditional electro-clinical classification of common familial epilepsies

The Epi4K Consortium*

*A list of authors, affiliations and contributions appears in Appendix 1.

Corresponding author:

Prof. Samuel F. Berkovic, M.D., F.R.S.

Epilepsy Research Centre, L2 Melbourne Brain Centre, 245 Burgundy Street

Heidelberg, VIC 3084, Australia.

Tel: +61 9035 7093

Fax: +61 9496 2291

Email: epi4k-fam@unimelb.edu.au

Keywords: epilepsy; phenotype; latent class analysis; genetics.

Text pages: 14

Words: 3,975

References: 32

Figures: 2

Tables: 4

Summary

Objective: Classification of epilepsy into types and subtypes is important for both clinical care and research into underlying disease mechanisms. A quantitative, data-driven approach may augment traditional electro-clinical classification and shed new light on existing classification frameworks.

Methods: We used latent class analysis, a statistical method that assigns subjects into groups called latent classes based on phenotypic elements, to classify individuals with common familial epilepsies from the Epi4K multiplex families study. Phenotypic elements included seizure types, seizure symptoms, and other elements of the medical history. We compared class assignments to traditional electro-clinical classifications and assessed familial aggregation of latent classes.

Results: A total of 1,120 subjects with epilepsy were assigned to 5 latent classes. Classes 1 and 2 contained subjects with generalized epilepsy, largely reflecting the distinction between absence epilepsies and younger onset (class 1) versus myoclonic epilepsies and older onset (class 2). Classes 3 and 4 contained subjects with focal epilepsies, and in contrast to classes 1 and 2, these did not adhere as closely to clinically defined focal epilepsy subtypes. Class 5 contained nearly all subjects with febrile seizures plus or unknown epilepsy type, as well as a few subjects with generalized epilepsy and a few with focal epilepsy. Family concordance of latent classes was similar to or greater than concordance of clinically defined epilepsy types.

Significance: Quantitative classification of epilepsy has the potential to augment traditional electro-clinical classification by (1) combining some syndromes into a single class, (2) splitting some syndromes into different classes, (3) helping to classify subjects who could not be classified clinically, and (4) defining the boundaries of clinically defined classifications. This approach can guide future research, including molecular genetic studies, by identifying homogeneous sets of individuals that may share underlying disease mechanisms.

Keywords: epilepsy; phenotype; latent class analysis; genetics.

Key Points

- Epilepsy phenotypes can be grouped based on phenotypic elements using a statistical classification method, latent class analysis.
- Latent classes preserved some traditional electro-clinical distinctions, e.g. generalized versus focal epilepsies.

- Some traditional syndromes were combined together or split apart, and some individuals were classified apart from their clinical syndromes.
- Family aggregation of latent classes was similar to or greater than aggregation of traditional phenotypes.
- Quantitative classification methods may help elucidate the shared and distinct biological mechanisms of different epilepsy phenotypes.

1. INTRODUCTION

Epilepsy is classified into clinically useful types and subtypes. Awareness of multiple different forms of epilepsy dates back to Hippocrates.[1] Modern efforts to codify the classification of epilepsies began with the first International League Against Epilepsy commission in the 1960s,[2] and classification continues to be refined and updated to this day.[3,4] Despite these efforts, the boundaries of classifications are often indistinct, and epilepsy classification remains challenging in some patients.

Classification of the epilepsies is important for clinical care and may guide us toward understanding the biology of these disorders. Accurate classification helps the clinician and the patient understand the natural history and prognosis of the disorder, informs the risk of comorbidities, and guides treatment decisions. Research into the underlying mechanisms of disease also requires accurate classification, as mechanisms often differ across subtypes of a disorder. For example, clinical genetic studies suggest both shared and distinct genetic determinants for different subtypes of epilepsy,[5–11] and identifying the underlying genetic determinants requires careful phenotyping and accurate classification into relatively homogeneous subgroups.

We sought to apply a quantitative, data-driven approach to the classification of epilepsy subtypes to augment traditional electro-clinical classification. Such approaches may combine existing syndromes into a single entity, separate existing syndromes into distinct subgroups, identify novel categories that were not clinically recognized, and reduce the subjectivity of classification by clinicians. Quantitative classification is particularly relevant to the study of common familial disorders, where genetic determinants are known to play a major role but are difficult to identify due in part to widespread phenotypic and genetic heterogeneity. Use of more data-driven phenotypic classes may help resolve such heterogeneity to improve gene discovery.

In this study, as a prespecified aim of the Epi4K Multiplex Families study,[12,13] we used latent class analysis to classify subjects with common familial epilepsies based on phenotypic elements. Latent class analysis is a statistical method that assigns subjects into subgroups, called latent classes, based on constellations of characteristics.[14] Compared to other methods of clustering analysis, this approach has the advantage of allowing for multiple correlated measurements and can therefore minimize the Type 1 error rate, improve statistical power, and eliminate the need to examine higher-order interactions in multivariate models. We analyzed the results of the latent class analysis by comparing these classifications to the epilepsy types assigned during clinical phenotyping in the Epi4K study, as well as the patterns of familial aggregation produced by each of these approaches.

2. METHODS

2.1 Ascertainment of families and data collection

Ascertainment methods are described in detail elsewhere[13] and summarized here briefly. Families were ascertained from 7 centers in North America, Europe, Australia and New Zealand. Families contained three or more relatives with unprovoked seizures of no known acquired cause. Data from every affected relative were obtained by a comprehensive protocol for data collection, assembling information from multiple sources including standardized diagnostic interviews with patients and relatives, collection and review of medical records, and systematic review of EEG and imaging reports. All of the assembled data were entered into a standardized diagnostic form, ensuring that key data elements were uniformly addressed for each family member with a history of seizures. Data from multiple sites were reviewed to ensure consistency of diagnostic methods across sites. The data were then synthesized by an expert clinician into electro-clinical diagnoses including seizure types and epilepsy syndromes. Potentially ambiguous seizure types (e.g. staring spells or convulsions) were classified as generalized or focal only when supported by EEG findings or a very compelling clinical history, otherwise they were considered unclassified seizure types. These determinations were made independently of other affected individuals in the same family.

Each individual was assigned to one of five epilepsy types: generalized, focal, combined (generalized and focal features in the same individual), febrile seizures plus (FS+), or unknown.

Within the generalized and focal epilepsy types, individuals were further assigned to subtypes, reflecting recognized epilepsy syndromes.[4]

For purposes of the current study, the main generalized epilepsy subtypes were as follows: absence epilepsy (including early-onset absence epilepsy (onset under four years), childhood absence epilepsy (onset 4 to 10 years) and juvenile absence epilepsy (onset older than 10 years)), juvenile myoclonic epilepsy (JME), generalized tonic-clonic seizures alone (GTCSA). As previously described,[13] we also had a category of "severe generalized epilepsy" (including epilepsy with myoclonic atonic seizures, absence epilepsy with eyelid myoclonia, epilepsy with myoclonic absences, and one case of Lennox-Gastaut syndrome of unknown cause) and a category of "other generalized" (including individuals with generalized epilepsy but without one of the above syndromes).

Focal epilepsy subtypes were as follows: temporal lobe epilepsy; frontal lobe epilepsy; posterior quadrant epilepsy (parietal, occipital, and posterior temporal regions); unknown localization epilepsy; and self-limited focal epilepsies of childhood (SLFE, including self-limited focal epilepsy with centrotemporal spikes and self-limited focal occipital epilepsies).

Finally, each family was classified as generalized, focal, mixed, or genetic epilepsy with febrile seizures plus (GEFS+), as described previously.[13]

2.2 Latent class analysis: variable selection

Fifteen variables were selected for inclusion in the latent class models.

- Four historical features: age of onset, history of febrile seizure, circadian pattern of seizure occurrence, and number of unprovoked seizures (of any type).
- Six seizure types: absence, myoclonic, generalized tonic-clonic (GTC), focal aware seizures (FAS), focal impaired-awareness seizures (FIAS), and focal to bilateral tonic-clonic seizures (focal-BTC).
- Five focal seizure symptoms: motor, sensory, psychic, autonomic, and aphasia.

Age of onset was defined as age at first unprovoked seizure (excluding febrile or provoked seizures) and for this analysis was dichotomized to above or below the median value of 10 years. Circadian pattern had three levels: (i) seizures predominantly while awake, (ii) seizures predominantly while asleep, and (iii) seizures during both waking and sleep or unknown. Number of unprovoked seizures had three levels: (i) 1-2 unprovoked seizures, (ii) >2 unprovoked

seizures, (iii) unknown number of seizures. All other variables were coded as either present or absent.

These variables were selected for inclusion in the latent class models because they are among the clinical features that underlie the basis for traditional electro-clinical classification. MRI data were not included in our models because our inclusion criteria specified persons with non-lesional epilepsies.

EEG data were not included because we relied heavily on EEG findings for classification of seizure types, especially the distinction between generalized and focal onset seizures. The final determination that a subject had generalized seizure types required evidence of generalized epileptiform EEG abnormalities except in exceptional cases where clinical symptoms unambiguously suggested generalized seizures. Similarly, the presence of focal epileptiform abnormalities was used to confirm the diagnosis of focal seizure types. Since seizure classification was based on EEG findings, and seizure types formed a large part of our latent class analysis, adding EEG results to our models would not add additional information.

Comorbidities such as intellectual disability, and clinical outcomes such as pharmacoresistance, were not included because these features, while associated with some epilepsy syndromes, are not the basis for the epilepsy classifications in this familial cohort. Additionally, moderate or severe intellectual disability was an exclusion criterion for our cohort of familial epilepsies.[13]

2.3 Latent class analysis: statistical methods

We conducted latent class analysis using the R package *LCAextend*,[15,16] which allows for categorical phenotypic elements and can also handle missing data. We used Bayesian Information Criterion to select the number of latent classes that best fit the data. We assessed associations between class assignments and specific input variables using chi-square tests, adjusting for multiple comparisons using a Bonferroni correction. Analyses were performed in the R programming language.

We assessed aggregation of class assignments within families in two ways. We first determined the proportion of families concordant for epilepsy type and concordant for latent class (see Results for definitions). Next, to allow for families showing evidence of familial aggregation for more than 1 class, we also constructed Krippendorf's alpha coefficient[17] using

the R package *irr*. To determine whether classes clustered within families more often than expected by chance, we compared the observed concordance rates and Krippendorf's alpha coefficients to empirical null distributions generated using permutation procedures under the assumption of no clustering.[18] We constructed nonparametric bootstrap confidence intervals for both the family concordance rates and Krippendorf's alpha coefficients using the *R* library *boot*.

3. RESULTS

The cohort included 1,120 individuals with epilepsy from 303 families. The distribution of epilepsy types in individual subjects was as follows: generalized 510, focal 321, combined 63, FS+ 27, unknown 199. Additional subject characteristics are described elsewhere in detail.[13]

Latent class analysis yielded 5 classes. In the models overall, each variable significantly contributed to the classifications ($p < 0.001$) suggesting that no variable was redundant.

3.1 Comparing latent classes to epilepsy types and subtypes

Latent class assignments are compared to the epilepsy types and subtypes assigned during clinical phenotyping in **Figure 1**. Latent class assignments preserved the distinction between generalized epilepsy and focal epilepsy but split each of these epilepsy types into two classes. Individuals with combined epilepsy were assigned mostly to class 4, with the remainder spread across three other classes. Individuals with the epilepsy types FS+ and unknown were classed together and made up most of class 5.

Individuals with generalized epilepsy were split into classes 1 and 2. The variables associated with these class assignments are shown in **Table 1** and the distribution of clinically designated generalized epilepsy subtypes in classes 1 and 2 is shown in **Table 2**. The variables associated with class 1 are characteristic of the absence epilepsies and almost all clinically designated absence epilepsy cases were in Class 1. Class 2 was associated with variables characteristic of JME and GTCSA and this was reflected in the designation of clinically defined cases. Individuals with "severe generalized" epilepsy were all assigned to class 1. The syndromes included in this category (see Methods) tend to occur at young ages and include absence seizures as a prominent seizure type. Indeed, all of these 23 subjects had ages of onset below the median of 10 years, and 20/23 had absence seizures. Notably, many of the subjects with "severe generalized" epilepsy also had myoclonic (14/23) and/or GTC (15/23) seizures,

variables associated with class 2, but given less weight than age of onset and absence seizures by the model. A similar pattern was seen in individuals with "other generalized" epilepsy, who were divided among classes 1 and 2: those subjects assigned to class 1 had younger onset and were more likely to have absence seizures (each comparison $p < 0.001$), while myoclonic and GTC seizures did not significantly contribute to class assignments in this subgroup.

Individuals with focal epilepsy were split into classes 3 and 4. The variables associated with these class assignments are shown in **Table 3**, and the distribution of focal epilepsy subtypes in classes 3 and 4 is shown in **Table 4**. Frontal lobe epilepsy and self-limited focal epilepsies (SLFE) were grouped together in class 4. The majority of individuals with temporal lobe epilepsy were assigned to class 3, but one-third were in class 4. In an analysis limited to individuals with temporal lobe epilepsy, those assigned to class 4 were younger, more likely to have motor symptoms and focal impaired-awareness seizures, and less likely to have sensory symptoms, psychic symptoms, or focal aware seizures than those assigned to class 3 (each comparison $p < 0.001$). That is, they resembled other subjects assigned to class 4 across a range of variables, rather than being assigned to class 4 based on a single powerful variable. Review of these cases suggested that most were diagnosed with temporal lobe epilepsy based on EEG data, which was not included in our latent class models. Posterior quadrant and unlocalized focal epilepsy were each split across classes 3 and 4, suggesting the model did not recognize these as coherent sets of individuals. Among individuals with posterior quadrant epilepsy in whom more specific localization was available (e.g. parietal lobe, occipital lobe) class assignments did not correspond to these localizations.

Individuals with combined generalized and focal epilepsy, a recently recognized ILAE epilepsy type,[4] were assigned to four different classes, although the majority (44/63, 70%) were assigned to class 4. In analysis limited to these subjects, the variables associated with assignment to class 4 versus any other class were the presence of motor symptoms ($p = 0.001$) and focal-BTC seizures ($p = 0.003$), and, less strongly, the presence of FIAS ($p = 0.03$) and sensory symptoms ($p = 0.05$). These findings were notable because focal-BTC seizures were not significantly associated with class 4 in subjects with focal epilepsy, but did seem to play a role in the classification of those with combined epilepsy. Similarly, in subjects with focal epilepsy, the presence of sensory symptoms was associated with class 3 rather than class 4 membership, but in those with combined epilepsy this pattern of association was reversed.

Class 5 contained nearly all individuals with unknown (195/199, 98%) and FS+ (26/27, 96%) epilepsy types. These individuals generally lacked all of the seizure types and symptoms considered in our analysis, which is why their epilepsy type was Unknown during clinical phenotyping. Class 5 also contained a small number of individuals with generalized (n=16) and focal (n=10) epilepsy types. The reason for this was not immediately clear. It was not because these individuals lack definable seizure types: 15/16 with generalized epilepsy had a generalized seizure type (GTC seizures) and 9/10 with focal epilepsy had at least one focal seizure type or seizure symptom.

These "outlier" individuals are of interest because they may reveal constellations of phenotypic features that make an individual too atypical to be grouped with other generalized or focal epilepsies, shedding light on the boundaries of those classifications. The pattern of seizure types and seizure symptoms in these individuals is shown in **Supplementary Table 1**. Of the 16 subjects with generalized epilepsy assigned to class 5, the pattern most predictive of assignment to class 5 was lack of absence or myoclonic seizure types, and presence of GTC seizures primarily during sleep. Of the 10 subjects with focal epilepsy assigned to class 5, lack of motor, sensory or psychic seizure symptoms along with lack of focal-BTC seizure types was a strongly predictive pattern. These patterns do not explain all of the outliers, but they demonstrate the potential for quantitative classification to reveal constellations of phenotypic elements that may lie beyond the boundaries of a particular group. Presumably, the other outlier individuals were similarly assigned to class 5 based on interactions among multiple variables that we were not able to deconstruct.

3.2 Family concordance of latent classes

Our previous report of this cohort demonstrated aggregation of epilepsy types within families (generalized versus focal epilepsy families), as well as familial aggregation of some epilepsy subtypes (absence epilepsies versus juvenile myoclonic epilepsy).[13] In that study, we defined concordant families as those where every individual with a definable epilepsy type had the same type; other individuals could have Unknown epilepsy type, but not a different definable epilepsy type.[13] In this study, to allow comparisons of concordance for epilepsy types versus latent classes, we assessed concordance of latent classes using analogous criteria, treating Class 5 as the equivalent of Unknown epilepsy type. If every individual assigned to classes 1-4 shared the

same class, the family was coded as concordant. Other relatives could be assigned to class 5, but not to a discordant class 1-4 (**Figure 2**). This procedure applied only to families with two or more members assigned a class other than 5; a family with every member belonging to class 5 was not considered concordant, because in our previous study of epilepsy types families could not consist entirely of individuals with epilepsy type Unknown.

A total of 138/303 (46%) families were concordant for latent classes. This included 68 families concordant for class 1, 20 families concordant for class 2, 17 families concordant for class 3, and 33 families concordant for class 4. Permutation analysis confirmed that each of these concordance frequencies is greater than expected by chance (each $p < 0.001$). In total, the number of families concordant for class 1 plus class 2 (88/303, 29%; 95% C.I. 24-34%) was greater than the number of families concordant for any generalized epilepsy subtype (59/303 families, 19%; 95% C.I. 16-23%). The total number of families concordant for class 3 plus class 4 (50/303 families, 17%; 95% C.I. 13-21%) was not statistically different from the number of families concordant for any focal epilepsy subtype (36/303 families, 12%; 95% C.I. 9-14%).

An alternative measure of familial aggregation is Krippendorf's alpha, which reflects concordance of class assignments among family members, with possible values ranging from 0 (no concordance) to 1 (perfect concordance). This analysis showed significant familial aggregation of latent classes (Krippendorf's alpha = 0.43; 95% CI 0.37, 0.49; $p < 0.001$) compared to the null hypothesis of no familial aggregation. These results were similar to the familial aggregation of epilepsy types (Krippendorf's alpha = 0.48; 95% CI 0.42, 0.54; $p < 0.001$).

4. DISCUSSION

This study used latent class analysis to augment classification in a large cohort of subjects with common epilepsies based only on their phenotypic elements, independent of their clinical syndrome classifications, but incorporating traditional electro-clinical seizure data. The resulting class assignments were broadly congruent with classical clinical classification, such as distinctions between generalized and focal epilepsies and between absence versus myoclonic subtypes of generalized epilepsy. Class assignments also combined some epilepsy subtypes and separated others in ways that reveal new insights about these categories and may facilitate the

search for genetic determinants. Family concordance for latent classes was similar or greater for latent classes than for epilepsy subtypes.

Several aspects of the results warrant highlighting and discussion. Classes 1 and 2 mapped closely to the absence epilepsies and juvenile myoclonic epilepsy, respectively, confirming the existing clinical framework for classification of patients with these generalized epilepsy syndromes. On the other hand, individuals who were clinically diagnosed with generalized epilepsy with only generalized tonic-clonic seizures primarily during sleep were identified by the model as "outliers" who were assigned to class 5. Similarly, four individuals with absence epilepsies were assigned to class 2, while nine individuals with JME or GTCSA were assigned to class 1. These classifications seemed to be driven largely by age of onset, supporting the clinical framework that age of onset is a defining feature of different generalized epilepsy subtypes and syndromes. Identifying these "outliers" helps define the boundaries of diagnostic categories and may inform genetic association studies, which may benefit from excluding these outlier cases.

Classes 3 and 4 contained individuals with focal epilepsies and mapped roughly to temporal lobe epilepsy and extratemporal epilepsy, respectively. Variables that can be seen in both temporal and extratemporal seizures, such as autonomic symptoms and focal to bilateral tonic-clonic seizures, were not distinguished by these class assignments. Seizures during sleep were associated with class 4, consistent with evidence that seizures during sleep are more common in frontal lobe epilepsy than in temporal lobe epilepsy.[19–21] The model grouped together frontal lobe epilepsy with self-limited focal epilepsies of childhood; these epilepsy subtypes have many features in common (nocturnal seizures, motor symptoms) and although they present as distinct clinical entities, their similarities may hint at shared underlying mechanisms. Subjects with posterior quadrant epilepsies were divided among the two classes. Seizures originating from the posterior quadrant often produce nonspecific symptoms and can be difficult to classify. The splitting of these individuals into two classes may be seen as a limitation of our model's ability to identify this group; alternatively, this may in fact be a more heterogeneous group than other types of focal epilepsy, with various genetic determinants.

We chose to perform this analysis in a cohort of familial epilepsies because we are particularly interested in the potential of quantitative classification methods to aid the discovery of the genetic determinants of epilepsy by helping to define phenotypically (and possibly

genetically) homogeneous sets of individuals. To that end, we should expect class assignments to aggregate within families, as family members with similar phenotypes are likely to share genetic determinants and should be grouped together for genetic analysis. In our study, familial aggregation of latent classes overall was similar to or greater than aggregation of clinically defined epilepsy types. Importantly, our latent class model was naïve to family relationships, treating each subject independently when assigning latent classes. This suggests that latent classes may be detecting patterns among subjects that are relevant to the familial nature of their epilepsies, and may be useful groups in which to search for shared genetic determinants.

To our knowledge, this is the first attempt at quantitative classification of broad epilepsy phenotypes. Studies with similar objectives have been performed in other fields, such as ADHD[22] and autism[23] where quantitative classification helped identify a region of interest for susceptibility genes. Within epilepsy, quantitative models have been used for more narrowly defined classifications, such as cognitive phenotypes,[24] depression phenotypes[25] and medication adherence.[26] Quantitative models have been used to study the overlap between autism and epilepsy,[27] and to classify different subtypes of psychogenic nonepileptic events.[28] Multivariate models of phenotypic elements have been used to predict seizure recurrence after first seizure and after withdrawal of antiepileptic medication.[29,30] However, the usefulness of this approach in classifying broad epilepsy phenotypes has not previously been explored.

Several of the potential advantages of quantitative classification are demonstrated here. First, some epilepsy subtypes were combined together (e.g. absence epilepsies and the "severe" generalized epilepsies), emphasizing their common features and suggesting the possibility of common underlying mechanisms. This finding reinforces molecular genetic discoveries showing the genes for severe generalized epilepsies are also relevant to the common GGE, such as *SLC2A1* in glucose transporter 1 deficiency syndrome.[31,32] Second, some epilepsy subtypes were separated into two different classes (e.g. temporal lobe epilepsy, posterior quadrant epilepsy), suggesting more heterogeneous groups that may benefit from further sub-classification and may have distinct underlying mechanisms. Third, some individuals were identified as atypical for a particular group, such as those with generalized tonic-clonic seizures during sleep who were not classified with other generalized epilepsies. Finally, quantitative classification helped classify some individuals who could not be classified clinically. Although our model did not achieve this in patients with Unknown epilepsy type, it did sort individuals with "other generalized" and

"unlocalized focal" epilepsies into meaningful classes. All of these findings are steps toward accurately defining homogeneous groups, which is essential for understanding the underlying mechanisms and determinants, as well as informing clinical management.

Our study has several limitations. The dataset was a single cohort of individuals with familial epilepsies. Latent class analyses are inherently dependent on the input variables and the composition of the cohort. Cohorts with different subject characteristics or different collections of phenotypic variables may produce different patterns of latent classes, and future studies should replicate this methodology in different cohorts. Future studies also should validate the biological relevance of quantitative approaches to epilepsy classification, an important hypothesis generated by our study. Some input variables in our model are based on clinical interpretation (e.g. distinguishing an absence seizure from a focal impaired-awareness seizure) and reflect the clinician's knowledge of the traditional electro-clinical classification paradigm. In general, the more objective data (free from clinical interpretation) goes into a model, the less it may recapitulate existing classification schemes. We used objective data (especially EEG results) to inform our classification of seizure types, and so these data were not included separately in our models. Other clinically important variables were not available for inclusion in our models, such as intellectual disability (an exclusion criterion in our familial cohort), and drug-resistance. The contributions of these variables should be explored by future studies.

In conclusion, we find that latent class analysis is a potentially valuable tool to classify subjects with common familial epilepsies and reveals new insights about the relationships among different epilepsy phenotypes. This approach may be useful to guide and inform studies of underlying mechanisms, including genetic determinants, which depend on the accurate identification of homogeneous sets of individuals likely to share common mechanisms.

ACKNOWLEDGEMENTS

We thank the families for participating in this study. This project was supported by an NINDS National Institute of Health grant (ID: U01NS077367). S.F.B. and I.E.S. were supported by an Australian National Health and Medical Research Council Program Grant (ID: 628952) and Practitioner Fellowship (I.E.S). R.O. was supported by NIH grants R01 NS078419, R01 NS104076 and RM1 HG007257. M.P.E. was supported by NIH grant R01 GM117946. M.I.R., W.O.P., R.H.T. and P.E.M. were supported by the National Institute of Social Care and Health Research, Epilepsy Research UK and the Waterloo Foundation. L.G.S and I.E.S were supported

by a New Zealand HRC grant (10/402) and Curekids. C.A.E. was supported by a Ruth L. Kirschstein National Research Service Award (NRSA) Institutional Research Training Grant, T32 NS091008-01.

DISCLOSURE

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

REFERENCES

1. Wolf P. History of epilepsy: nosological concepts and classification. Epileptic Disord. 2014; 16(3):261–9.
2. Gastaut H, Caveness W, Landolt H, et al. A proposed international classification of epileptic seizures. Epilepsia. 1964; 5:297–306.
3. Fisher RS, Cross JH, French JA, et al. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017; 58(4):522–30.
4. 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):512–21.
5. Berkovic SF, Howell RA, Hay DA, et al. Epilepsies in twins: Genetics of the major epilepsy syndromes. Ann Neurol. 1998; 43(4):435–45.
6. Ottman R, Lee JH, Hauser WA, et al. Are generalized and localization-related epilepsies genetically distinct? Arch Neurol. 1998; 55(3):339–44.
7. Winawer MR, Rabinowitz D, Pedley TA, et al. Genetic influences on myoclonic and absence seizures. Neurology. 2003; 61(11):1576–81.
8. Marini C, Scheffer IE, Crossland KM, et al. Genetic architecture of idiopathic generalized epilepsy: clinical genetic analysis of 55 multiplex families. Epilepsia. 2004; 45(5):467–78.
9. Winawer MR, Marini C, Grinton BE, et al. Familial clustering of seizure types within the idiopathic generalized epilepsies. Neurology. 2005; 65(4):523–8.
10. Peljto AL, Barker-Cummings C, Vasoli VM, et al. Familial risk of epilepsy: a population-

based study. Brain. 2014; 137(3):795–805.

11. Tobochnik S, Fahlstrom R, Shain C, et al. Familial aggregation of focal seizure semiology in the Epilepsy Phenome/Genome Project. Neurology. 2017; 89(1):22–8.
12. Epi4K Consortium. Epi4K: Gene discovery in 4,000 genomes. Epilepsia. 2012; 53(8):1457–67.
13. Epi4K Consortium. Phenotypic analysis of 303 multiplex families with common epilepsies. Brain. 2017; 140(8):2144–56.
14. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci. 2013; 14(2):157–68.
15. Labbe A, Bureau A, Merette C. Integration of genetic familial dependence structure in latent class models. Int J Biostat. 2009; 5(1):Article 6.
16. Tayeb A, Labbe A, Bureau A, et al. Solving genetic heterogeneity in extended families by identifying sub-types of complex diseases. Comput Stat. 2011; 26:539–60.
17. Krippendorf K. Content analysis: an introduction to its methodology. Beverly Hills: Sage Publications, Inc.; 1980.
18. Winawer M, Ottman R, Rabinowitz D. Concordance of disease form in kindreds ascertained through affected individuals. Stat Med. 2002; 21(13):1887–97.
19. Crespel A, Coubes P, Baldy-Moulinier M. Sleep influence on seizures and epilepsy effects on sleep in partial frontal and temporal lobe epilepsies. Clin Neurophysiol. 2000; 111 Suppl 2:S54-9.
20. Derry CP. The Sleep Manifestations of Frontal Lobe Epilepsy. Curr Neurol Neurosci Rep. 2011; 11(2):218–26.
21. Winawer MR, Shih J, Beck ES, et al. Genetic effects on sleep/wake variation of seizures. Epilepsia. 2016; 57(4):557–65.
22. Rasmussen ER, Neuman RJ, Heath AC, et al. Familial clustering of latent class and DSM-IV defined attention-deficit/hyperactivity disorder (ADHD) subtypes. J Child Psychol Psychiatry. 2004; 45(3):589–98.
23. Shao Y, Cuccaro ML, Hauser ER, et al. Fine mapping of autistic disorder to chromosome 15q11-q13 by use of phenotypic subtypes. Am J Hum Genet. 2003; 72(3):539–48.
24. Hermann BP, Zhao Q, Jackson DC, et al. Cognitive phenotypes in childhood idiopathic epilepsies. Epilepsy Behav. 2016; 61:269–74.

25. Rayner G, Jackson GD, Wilson SJ. Two distinct symptom-based phenotypes of depression in epilepsy yield specific clinical and etiological insights. Epilepsy Behav. 2016; 64(Pt B):336–44.
26. Modi AC, Rausch JR, Glauser TA. Patterns of nonadherence to antiepileptic drug therapy in children with newly diagnosed epilepsy. JAMA. 2011; 305(16):1669–76.
27. Cuccaro ML, Tuchman RF, Hamilton KL, et al. Exploring the relationship between autism spectrum disorder and epilepsy using latent class cluster analysis. J Autism Dev Disord. 2012; 42(8):1630–41.
28. Bodde NMG, van der Kruijs SJM, Ijff DM, et al. Subgroup classification in patients with psychogenic non-epileptic seizures. Epilepsy Behav. 2013; 26(3):279–89.
29. Bonnett LJ, Marson AG, Johnson A, et al. External validation of a prognostic model for seizure recurrence following a first unprovoked seizure and implications for driving. Brock G, editor. PLoS One. 2014; 9(6):e99063.
30. Lamberink HJ, Otte WM, Geerts AT, et al. Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. Lancet Neurol. 2017; 16(7):523–31.
31. Arsov T, Mullen SA, Rogers S, et al. Glucose transporter 1 deficiency in the idiopathic generalized epilepsies. Ann Neurol. 2012; 72(5):807–15.
32. Epi4K Consortium, Epilepsy Phenome/Genome Project. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. Lancet Neurol. 2017; 16(2):135–43.

FIGURE LEGENDS

Figure 1. Comparison of latent class assignments versus epilepsy types assigned by clinical phenotyping. Abbreviations: FS+, febrile seizures plus.

Figure 2. Examples of family concordance for different classification methods. Concordant and discordant families are shown for epilepsy types (colors), epilepsy subtypes (first line below each pedigree symbol), and latent classes (second line below each pedigree symbol). Concordant

families could include individuals with unknown epilepsy type/subtype, or latent class 5, but not individuals with from other discordant categories.

Abbreviations: FLE, frontal lobe epilepsy; JME, juvenile myoclonic epilepsy; TLE, temporal lobe epilepsy; SLFE, self-limited focal epilepsy.

Table 1. Variables associated with sorting individuals with generalized epilepsy into class 1 versus class 2.

Variable	Class 1	Class 2	*p*(corr)[a]
Age	younger	older	<0.001
Absence	+++	---	<0.001
Myoclonic	---	+++	<0.001
GTC	---	+++	<0.001
Febrile	+	-	0.007
Circadian			ns
Number			ns

[a]Chi-square test with Bonferroni correction for multiple comparisons.

Age of onset is relative to the median value of 10 years. The following variables included in the overall model pertain only to focal epilepsy, were not present in any individuals with generalized epilepsy and hence were not included in this analysis: focal aware seizures, focal impaired awareneness seizures, focal to bilateral tonic-clonic seizures, psychic seizures, sensory seizures, focal motor seizures, autonomic seizures, aphasic seizures.

Abbreviations: GTC, generalized tonic-clonic.

Table 2. Distribution of generalized epilepsy subtypes into class 1 versus class 2.

Clinical Classification	n of subjects	Class 1 n (%)	Class 2 n (%)
Absence	260	**256 (98%)**	4 (2%)
JME/GTCSA	141	9 (6%)	**132 (94%)**
Severe Generalized	23	**23 (100%)**	0 (0%)
Other Generalized	69	**46 (67%)**	23 (33%)

For each subtype the larger percentage value is bolded for ease of visual comparison.
Abbreviations: GTCSA, generalized tonic-clonic seizures alone; JME, juvenile myoclonic epilepsy.

Table 3. Variables associated with classifying individuals with focal epilepsy into class 3 vs class 4.

Variable	Class 3	Class 4	*p*(corr)[a]
Psychic	+++	---	<0.001
Sensory	+++	---	<0.001
FAS	+++	---	<0.001
Motor	---	+++	<0.001
FIAS	---	+++	<0.001
Age	older	younger	<0.001
Circadian	wake	sleep	<0.001
Febrile			ns
Autonomic			ns
Number			ns
Focal-BTC			ns
Aphasia			ns

[a]Chi-square test with Bonferroni correction for multiple comparisons.

Age of onset is relative to the median value of 10 years. The following variables included in the overall model pertain only to generalized epilepsy, were not present in any individuals with focal epilepsy and hence were not included in this analysis: absence seizures, myoclonic seizures, generalized tonic-clonic seizures.

Abbreviations: BTC, bilateral tonic-clonic; FAS, focal aware seizures; FIAS, focal impaired awareness seizures.

Table 4. Distribution of focal epilepsy subtypes into class 3 versus class 4.

Clinical Classification	n of subjects	Class 3 n (%)	Class 4 n (%)
Temporal	131	**89 (68%)**	42 (32%)
Frontal	25	2 (8%)	**23 (92%)**
SLFE	41	4 (10%)	**37 (90%)**
Posterior	64	32 (50%)	32 (50%)
Unlocalized	50	13 (26%)	**37 (74%)**

For each subtype the larger percentage value is bolded for ease of visual comparison.

Abbreviations: SLFE, self-limited focal epilepsy.

APPENDIX 1: Authors, contributions and affiliations

Core Working Group

CAE (primary data analysis), MPE (latent class derivations), SFB, RO, STB.

Drafting of Manuscript

CAE (wrote first draft), MPE, SFB, RO. All other authors reviewed, edited and approved the manuscript.

Epi4K Project Design and Steering Committee

DHL, DBG, SFB, ASA, PCo, DD, MPE, ELH, RKu, AGM, HCM, TJO, RO, StP, SlP, AP, IES.

Case Ascertainment and Phenotyping

University of Melbourne – SFB, IES, STB, SAM, DEC, RB, RHT.

University of Otago – LGS, SPa

Tel Aviv University – SK, ZA, HG, ADK.

Columbia University – RO, MRW, RL.

Swansea University – MIR, WOP, RHT, PEMS.

Université de Montréal – PCo, MG.

Royal College of Surgeons, Dublin – ND, MM.

EPGP – DHL, SC, CF, KM, PW, EBG, AP, AV, TG, JFB, SH, GV, EV, EK, GC, JSi, JSh, RKu, OD, JBl, PM, MS, RKn, HEK, SG, JMP, BA, JuP, KH, NBF, LLT, JW, RS.

Epi4K Collaborators

Bassel Abou-Khalil[1], Zaid Afawi[2], Andrew S. Allen[3], Jocelyn F. Bautista[4], Susannah T. Bellows[5], Samuel F. Berkovic[5], Judith Bluvstein[6], Rosemary Burgess[5], Gregory Cascino[7], Patrick Cossette[8], Sabrina Cristofaro[6], Douglas E. Crompton[5], Norman Delanty[9], Orrin Devinsky[6], Dennis Dlugos[10], Colin A. Ellis[5,11], Michael P. Epstein[12], Nathan B. Fountain[13], Catharine Freyer[14], Eric B. Geller[15], Tracy Glauser[16], Simon Glynn[17], Hadassa Goldberg-Stern[18], David B. Goldstein[19], Micheline Gravel[8], Kevin Haas[1], Sheryl Haut[20], Erin L. Heinzen[19], Heidi E. Kirsch[14], Sara Kivity[18], Robert Knowlton[21], Amos D. Korczyn[2], Eric Kossoff[22], Ruben Kuzniecky[6], Rebecca Loeb[23], Daniel H. Lowenstein[14], Anthony G. Marson[24], Mark McCormack[25], Kevin McKenna[14], Heather C. Mefford[26], Paul Motika[27], Saul A. Mullen[5], Terence J O'Brien[28], Ruth Ottman[23], Juliann Paolicchi[1], Jack M. Parent[17], Sarah Paterson[29], Steven Petrou[30], Slavé Petrovski[5,19,28], William Owen Pickrell[31], Annapurna Poduri[32], Mark I. Rees[31], Lynette G. Sadleir[29], Ingrid E. Scheffer[5,33], Jerry Shih[34], Rani Singh[35], Joseph Sirven[36], Michael Smith[27], Phil E.M. Smith[37], Liu Lin Thio[38], Rhys H. Thomas[5,31,39], Anu Venkat[40], Eileen Vining[22], Gretchen Von Allmen[41], Judith Weisenberg[38], Peter Widdess-Walsh[15] and Melodie R. Winawer[23].

Affiliations

1. Department of Neurology, Vanderbilt University Medical Center, Nashville, Tennessee 37232 USA.

2. Sackler School of Medicine, Tel Aviv University, Ramat Aviv, Israel.

3. Center for Statistical Genetics and Genomics, Duke University School of Medicine, Durham, North Carolina 27708, USA.

4. Department of Neurology, Cleveland Clinic Lerner College of Medicine & Epilepsy Center of the Cleveland Clinic Neurological Institute, Cleveland, Ohio 44195 USA.

5. Epilepsy Research Centre, Department of Medicine, University of Melbourne (Austin Health), Heidelberg, Victoria 3084, Australia.

6. NYU Comprehensive Epilepsy Center, New York University, Department of Neurology, NYU School of Medicine, New York, New York 10016, USA.

7. Division of Epilepsy, Mayo Clinic, Rochester, Minnesota 55905 USA.

8. Centre of Excellence in Neuromics and CHUM Research Center, Université de Montréal, CHUM-Hôpital, Notre-Dame Montréal, Quebec H2L 4M1, Canada.

9. Department of Neurology, Beaumont Hospital and Royal College of Surgeons, Dublin 9, Ireland.

10. Department of Neurology and Pediatrics, The Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

11. Department of Neurology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

12. Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, 30322, USA.

13. FE Dreifuss Comprehensive Epilepsy Program, University of Virginia, Charlottesville, Virginia 22908 USA.

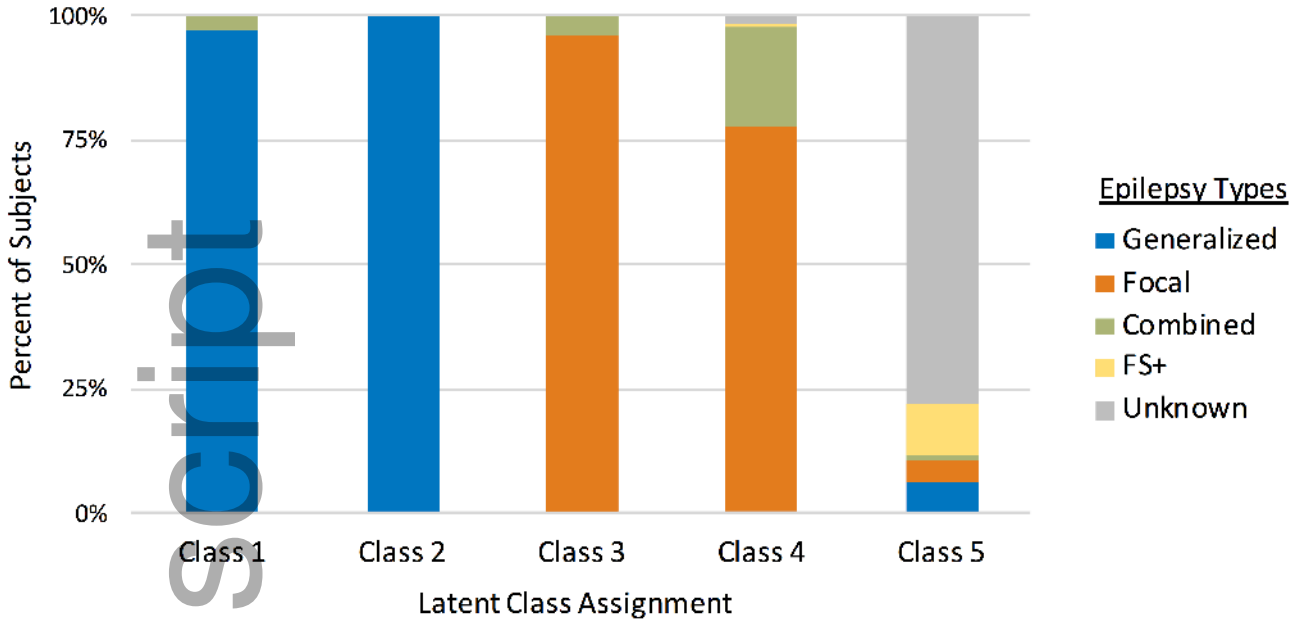
14. Department of Neurology, University of California, San Francisco, San Francisco, California 94143, USA.

15. Institute of Neurology and Neurosurgery at Saint Barnabas, Saint Barnabas Medical Center, Livingston NJ, USA.

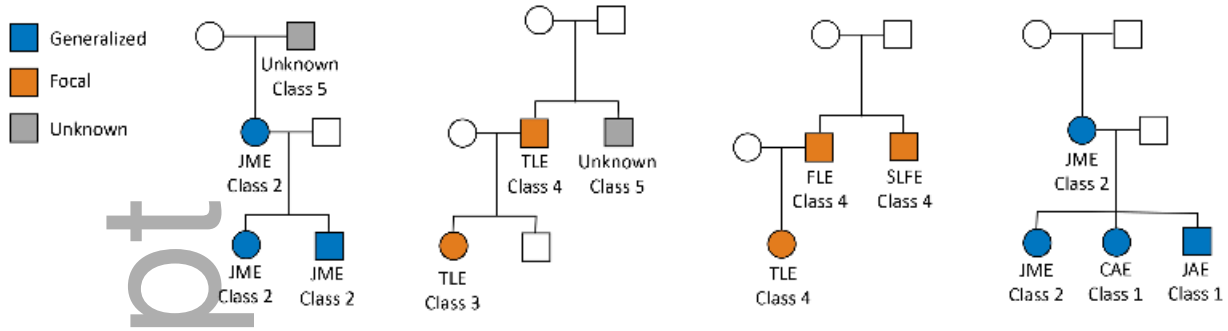
16. Division of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio

45229 USA.
17. Department of Neurology and Neuroscience Graduate Program, University of Michigan Medical Center, Ann Arbor, and Ann Arbor Veterans Administration Healthcare System, Ann Arbor, Michigan, USA.
18. Epilepsy Unit, Schneider Children's Medical Center of Israel, Petach Tikvah, Israel.
19. Institute for Genomic Medicine, Columbia University Irving Medical Center, New York, New
York 10032, USA.
20. Adult Epilepsy, Montefiore Einstein, 111 East 210th St, Bronx, NY 10467
21. University of Alabama at Birmingham School of Medicine, Birmingham, Atlanta, USA.
22. Department of Neurology, Johns Hopkins Hospital, Baltimore, Maryland 21287 USA.
23. Departments of Epidemiology and Neurology, and the G. H. Sergievsky Center, Columbia University; and Division of Translational Epidemiology, New York State Psychiatric Institute, New York, New York 10032, USA.
24. Department of Molecular and Clinical Pharmacology, University of Liverpool, Clinical Sciences Centre, Lower Lane, Liverpool L9 7LJ, UK.
25. Molecular & Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin 2, Ireland
26. Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, Washington 98195, USA.
27. Comprehensive Epilepsy Center, Oregon Health and Science University, Portland, OR 97239 USA.
28. Departments of Medicine and Neurology, The Royal Melbourne Hospital, Parkville, Victoria 3050, Australia.
29. Department of Paediatrics, School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand.
30. Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville, Victoria 3052, Australia
31. Wales Epilepsy Research Network, Swansea University Medical School, Swansea University, Wales, UK.

32. Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, and Department of Neurology, Harvard Medical School, Boston, Massachusetts 02115, USA.
33. Department of Paediatrics, Royal Children's Hospital, Melbourne, Victoria, Australia.
34. Comprehensive Epilepsy Center, Department of Neurosciences, University of California, San Diego School of Medicine, San Diego, California, USA.
35. Carolinas Health Care System, USA.
36. Department of Neurology, Mayo Clinic, Scottsdale, Arizona 85259 USA.
37. Department of Neurology, University Hospital of Wales, Heath Park, Cardiff, Wales, CF14 4XW, UK.
38. Department of Neurology, Washington University School of Medicine, St. Louis, Missouri 63110.
39. MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University, Wales, CF24 4HQ, UK.
40. Department of Pediatrics, Children's Hospital at Saint Peter's University Hospital, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA.
41. Division of Child & Adolescent Neurology, Departments of Pediatrics, University of Texas Medical School, Houston, Texas 77030 USA.



epi_16354_f1.tif



	Family A	Family B	Family C	Family D
Epilepsy type:	Concordant	Concordant	Concordant	Concordant
Epilepsy subtype:	Concordant	Concordant	Discordant	Discordant
Latent class:	Concordant	Discordant	Concordant	Discordant

epi_16354_f2.tif