RESEARCH ARTICLE

# Predicting functional effects of ion channel variants using new phenotypic machine learning methods

Christian Malte Boßelmann [1,2], Ulrike B. S. Hedrich[1], Holger Lerche[1☯*], Nico Pfeifer [2,3☯*]

1 Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany, 2 Methods in Medical Informatics, Department of Computer Science, University of Tuebingen, Tuebingen, Germany, 3 Institute for Bioinformatics and Medical Informatics, University of Tuebingen, Tuebingen, Germany

☯ These authors contributed equally to this work.
* holger.lerche@med.uni-tuebingen.de (HL); nico.pfeifer@uni-tuebingen.de (NP)

## Abstract

Missense variants in genes encoding ion channels are associated with a spectrum of severe diseases. Variant effects on biophysical function correlate with clinical features and can be categorized as gain- or loss-of-function. This information enables a timely diagnosis, facilitates precision therapy, and guides prognosis. Functional characterization presents a bottleneck in translational medicine. Machine learning models may be able to rapidly generate supporting evidence by predicting variant functional effects. Here, we describe a multi-task multi-kernel learning framework capable of harmonizing functional results and structural information with clinical phenotypes. This novel approach extends the human phenotype ontology towards kernel-based supervised machine learning. Our gain- or loss-of-function classifier achieves high performance (mean accuracy 0.853 SD 0.016, mean AU-ROC 0.912 SD 0.025), outperforming both conventional baseline and state-of-the-art methods. Performance is robust across different phenotypic similarity measures and largely insensitive to phenotypic noise or sparsity. Localized multi-kernel learning offered biological insight and interpretability by highlighting channels with implicit genotype-phenotype correlations or latent task similarity for downstream analysis.

## Author summary

Genetic variants can impact the function of voltage-gated ion channels, which are transmembrane channels that are important for the electrical signalling of neurons. Individuals affected by these variants may have severe neurological or cardiac disease. Knowing the functional effects of variants can help us improve the diagnosis and clinical care for these individuals, and may also enable precision medicine aimed at correcting channel function. Machine learning methods are capable of predicting these functional effects, but previous models have only included information from protein sequence and structure. Here, we have developed a novel algorithm that can integrate clinical information by using the

Human Phenotype Ontology (HPO), a standardized vocabulary of phenotypic features encountered in human disease. These different sources of information are used to train a multi-kernel support-vector machine. Our results suggest that this model is robust, interpretable, and more accurate than previous methods. Variant functional effect prediction engines will be useful tools for clinical geneticists, neurologists and biophysicists.

## Introduction

Variants in genes encoding voltage-gated sodium channels are associated with a broad spectrum of often severe neurodevelopmental, neuromuscular, and cardiac diseases. Complementary genetic and electrophysiological investigations have contributed towards a growing list of these ion channel disorders or 'channelopathies' [1]. For example, variants in sodium channel protein type 1 subunit alpha (*SCN1A*, Na$_V$1.1) are the second-most common finding in cohorts of individuals with monogenic epilepsy [2]. The majority of these variants are associated with Dravet syndrome, a severe developmental and epileptic encephalopathy (DEE) and the most common drug-resistant genetic epilepsy [3]. Likewise, variants in *SCN5A* are responsible for >75% of cases with genetically confirmed Brugada syndrome (BrS), an abnormality of cardiac conduction with a significant risk of sudden death in young individuals [4]. Lastly, variants in *SCN9A* are associated with paroxysmal extreme pain disorder (PEPD), a rare autosomal-dominant disorder of pain processing and autonomic function with neonatal to infantile onset, where individuals experience sudden attacks of excruciating pain with no response to conventional analgesic medication [5]. These examples serve to illustrate the high burden of disease and unmet medical need in the affected individuals.

Missense variants can have an impact on the biophysical function of the mutant ion channel. Amino acid substitutions may alter the physiochemical properties including polarity, residue weight and size, disulphide linkage, or hydrophobicity of residues that are located in domains or functional motifs that are critical to channel function [6–8]. This may affect every aspect of channel function including voltage sensing, expression level or membrane trafficking, all parameters of channel gating such as the voltage-dependence or time constants of activation and fast or slow inactivation, incomplete channel inactivation leading to persistent sodium current, or loss of ion selectivity [9–12]. In the majority of cases, the sum of the functional consequences of a variant can be categorized as an overall net gain-of-function (GOF) or loss-of-function (LOF) [13]. Less often, mixed effects on channel function, downstream effects on neuronal firing or network dysfunction, or different functional results depending on the expression system contribute towards a variant for which an overall GOF or LOF may be difficult to determine [14,15].

These functional effects of missense variants correlate with distinct clinical phenotypes, which have been described for almost all voltage-gated sodium channels. While *SCN1A*-LOF is associated with Dravet syndrome, *SCN1A*-GOF causes familial hemiplegic migraine type 3 (FHM3) or a spectrum of DEE including movement disorder and arthrogryposis (NDEEMA) [16,17]. In *SCN5A*, LOF by altered channel inactivation or non-conducting channels leads to Brugada syndrome type 1, while GOF by increased persistent inward sodium current and longer action potential duration results in familial long QT syndrome type 3 (LQT3) [18–21]. Perhaps most strikingly, *SCN9A*-LOF causes paroxysmal extreme pain disorder, yet *SCN9A*-GOF leads to congenital insensitivity to pain [22]. For *SCN2A* and *SCN8A*, GOF is associated with an early-onset epilepsy phenotype including self-limited (familial) infantile *epilepsy* (*SeLIE*) or DEE. Individuals with these variants respond well to treatment with sodium channel blockers

(SCBs). Conversely, *SCN2A*-LOF or *SCN8A*-LOF leads to a later seizure onset and unfavourable response to SCBs [23,24]. Evidently, individual phenotypic features contain valuable information to guide variant interpretation.

Knowledge of variant functional effects and genotype-phenotype correlations facilitates a timely diagnosis, informs genetic counselling, and offers therapeutic and prognostic guidance to the clinician [25]. However, variant interpretation requires causal evidence generated by electrophysiological experiments which are relatively time-consuming and expensive. Recent advances in next-generation sequencing techniques have generated a wealth of genetic variation which threatens to outpace the ability of experimental assays to keep up with the increased demand, resulting in a growing number of individuals affected by variants of unknown functional consequence [26]. Positioned at this interface between clinical genetics and electrophysiology, novel computational tools may be helpful to generate supporting evidence. Previously, Heyne et al. demonstrated the use of a machine learning method, gradient boosting, to predict the overall functional effect of missense variants in voltage-gated sodium or calcium channels [27]. They used known genotype-phenotype correlations to infer likely functional labels for supervised learning, which may not make full use of available clinical information and risks introducing bias and uncertainty to the training data set. In this study, we extend our own previous work on kernel-based multi-task learning methods applied to potassium channel variants [28] by designing a framework capable of integrating functional, structural and clinical data while leveraging prior knowledge on channel family structure.

## Results

### Learning with phenotypic similarity outperforms both conventional baseline and state-of-the-art models

We first trained naïve baseline models, one channel-specific support-vector machine (SVM) for variants from every single voltage-gated sodium channel (Dirac), and one global SVM for variants from all voltage-gated sodium channels (Union). We then implemented a method from our previous work: Multi-task learning (MTL) SVM embeds explicit prior knowledge on the structural and evolutionary pairwise similarity between ion channels in the feature space induced by our kernel function, which enables the model to assign more weight when learning from variants of similar channels [28]. MTL-SVM outperformed the baseline models across all metrics (Table 1). Finally, our novel multi-task multi-kernel learning model (MT)MKL approach for GOF/LOF prediction was trained with the Jaccard similarity measure and uniform kernel weights. Including phenotypic similarity information in the machine learning model further improves predictive performance (Table 1 and Fig 1). For comparison to current state-of-the-art methods, we trained a gradient boosting machine as used in previous work by Heyne et al. and the decision rule introduced by Brunklaus et al. [27,29]. For internal validation, these methods were trained and tested on the same data splits of our pre-processed training data using the k-fold nested cross-validation procedure described above. For external

**Table 1. Nested cross-validation (outer folds) model performance estimates for naïve baseline models (Dirac, Union), our previous state-of-the-art multi-task learning model (MTL), and phenotypic multi-task multi-kernel learning (MKL).** Metrics are reported as mean ± standard deviation (SD).

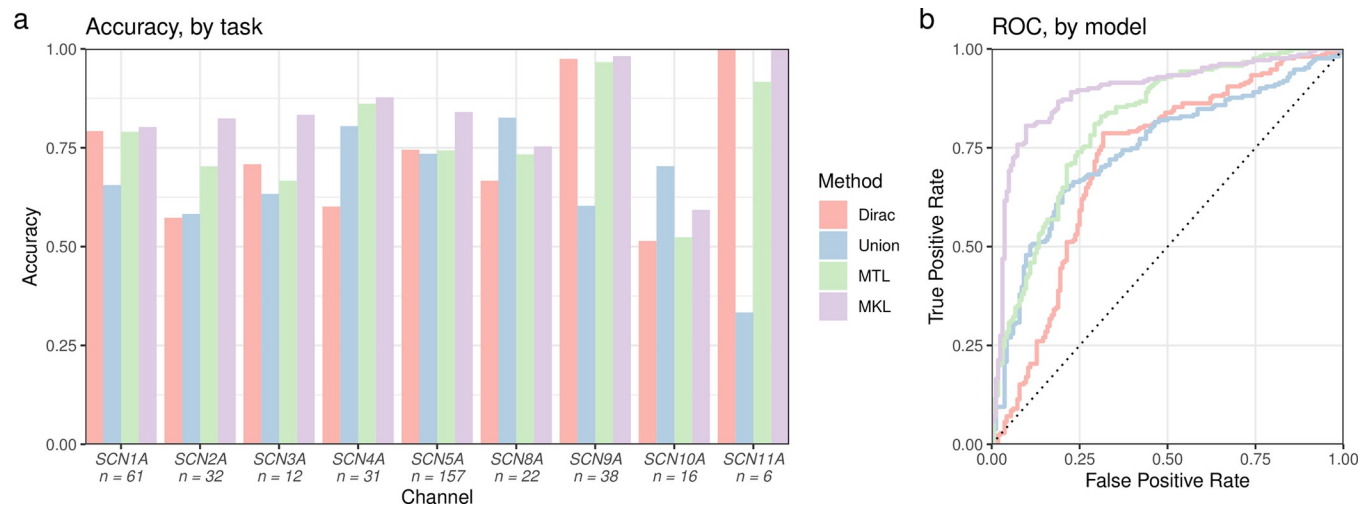| Kernel | ACC | $F_1$ score | Cohen's κ | MCC | AU-PRC | AU-ROC |
|---|---|---|---|---|---|---|
| Dirac | 0.741 ± 0.033 | 0.698 ± 0.038 | 0.472 ± 0.066 | 0.475 ± 0.065 | 0.694 ± 0.049 | 0.731 ± 0.036 |
| Union | 0.693 ± 0.063 | 0.649 ± 0.089 | 0.379 ± 0.127 | 0.382 ± 0.124 | 0.679 ± 0.154 | 0.761 ± 0.082 |
| MTL | 0.760 ± 0.025 | 0.698 ± 0.049 | 0.501 ± 0.059 | 0.508 ± 0.058 | 0.798 ± 0.040 | 0.818 ± 0.017 |
| MKL | **0.837 ± 0.035** | **0.824 ± 0.029** | **0.673 ± 0.067** | **0.677 ± 0.064** | **0.837 ± 0.039** | **0.890 ± 0.023** |

**Fig 1. Phenotypic multi-task multi-kernel learning outperforms both conventional and state-of-the-art models.** a: Cross-validation model performance estimates of mean accuracy for each task (channel). b: ROC curves. The AU-ROC for each model is shown in Table 1. Abbreviations: MTL–multi-task learning; MKL–multi-kernel learning; ROC–receiver operator characteristics.

validation, all three models were used to obtain predictions for a list of novel missense variants in voltage-gated sodium channels (S1 Table). Results are shown in Table 2. In both validation procedures, our multi-task multi-kernel learning model performed best.

## Learning with phenotypic similarity is robust across different similarity measures and largely insensitive to phenotypic noise or sparsity

We compared three commonly used similarity measures: Jaccard, Lin, and Resnik. The Jaccard similarity measure always results in a positive semi-definite kernel matrix, which extends naturally to kernel-based learning methods [30]. Measures based on information content (IC), such as Lin and Resnik, lead to indefinite kernel matrices for the human phenotype ontology, as it allows for multiple parent terms. This invalidates the underlying theory and would lead to a non-convex SVM optimization problem, which is computationally expensive to solve and is not guaranteed to find the global optimal solution. For similarity-based classification, three methods are available to approximate the nearest positive semi-definite similarity matrix to be

**Table 2. Performance estimates for other state-of-the-art models and phenotypic multi-task multi-kernel learning (MKL).** For internal validation, each model was assessed by nested k-fold cross-validation. For external validation on a set of novel missense variants, the web-based tools for the Gradient boosting and decision rule method were used (see 'Validation'). Note that AU-PRC and AU-ROC are not available for the decision rule as it is not a probabilistic classifier. Phenotypic learning outperforms other models across all metrics and is shown in bold. Abbreviations: ACC–accuracy; AU-PRC–area under the precision recall curve; AU-ROC–area under the receiver operator characteristics curve.

| Method | ACC | Sensitivity | Specificity | F₁ score | Cohen's κ | MCC | AU-PRC | AU-ROC |
|---|---|---|---|---|---|---|---|---|
| Internal validation | | | | | | | | |
| Decision rule | 0.136 ± 0.021 | 0.138 ± 0.019 | 0.712 ± 0.007 | 0.241 ± 0.031 | 0.073 ± 0.011 | 0.196 ± 0.015 | – | – |
| Gradient boosting | 0.659 ± 0.049 | 0.603 ± 0.059 | 0.702 ± 0.069 | 0.607 ± 0.051 | 0.306 ± 0.096 | 0.307 ± 0.097 | 0.639 ± 0.085 | 0.722 ± 0.061 |
| MKL | **0.837 ± 0.035** | **0.866 ± 0.033** | **0.814 ± 0.066** | **0.824 ± 0.029** | **0.673 ± 0.067** | **0.677 ± 0.064** | **0.837 ± 0.039** | **0.890 ± 0.023** |
| External validation | | | | | | | | |
| Decision rule | 0.033 | 0.025 | 0.672 | 0.048 | 0.000 | 0.000 | – | – |
| Gradient boosting | 0.667 | 0.500 | 0.750 | 0.500 | 0.250 | 0.250 | 0.730 | 0.595 |
| MKL | **0.967** | **0.900** | **1.000** | **0.947** | **0.923** | **0.926** | **1.000** | **1.000** |
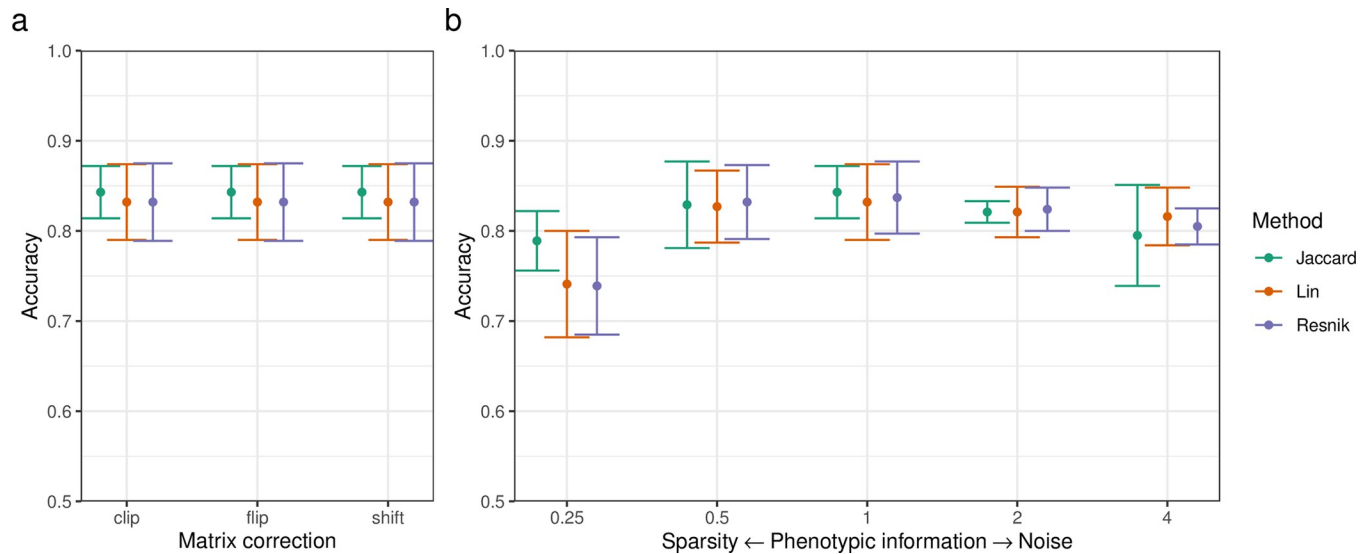
**Fig 2. Phenotypic multi-task multi-kernel learning performs well across different similarity measures or matrix correction methods and is largely insensitive to sparse or noisy phenotypic information.** *a*: Dot chart of MKL model mean accuracy and standard deviation for each similarity measure and methods for approximating the nearest PSD matrix. The Jaccard index does not require matrix correction but is shown for comparison. *b*: Dot chart of MKL model mean accuracy and standard deviation for each similarity measure across levels of simulated phenotypic sparsity and noise. Noise is shown as human phenotype ontology term ratio (e.g. noise ratio 0.25 removes 3 out of 4 terms from the set, while noise ratio 4 adds 3 random terms for each original term in the set).

used as kernel matrix: spectrum clip, flip and shift [31]. The choice of matrix correction method did not have an impact on MKL model performance (Fig 2A and Table 3).

Information content measures are insensitive to varying taxonomic link distances and adapt the static knowledge structure of an ontology to different contexts by including probabilistic estimates [32]. We hypothesized that these measures may perform better in imperfectly phenotyped cases. Traditional phenotyping is expensive and prone to misclassification error. Recent studies have introduced 'silver standard' phenotyping from data mining of electronic health records and demonstrated advances in automated phenotyping pipelines [33,34]. We simulated sparse phenotypes by removing terms from the set of all terms for each variant at random before computing semantic similarity. As more terms were removed, the information from the phenotype kernel matrix was reduced and the mean accuracy of the MKL model approached that of the MTL model (Fig 2B and Table 4). Likewise, we simulated noisy phenotypes by sampling (without replacement) additional terms from the set of all possible terms and adding these to the set of terms for each variant before computing semantic similarity. We observed that all semantic similarity measures resulted in models with a stable accuracy even at high levels of simulated noise (Fig 2B and Table 4).

**Table 3. Nested cross-validation (outer folds) model performance estimates.** Model performance of semantic similarity measures over methods of approximating the nearest PSD matrix. The Jaccard index does not require matrix correction but is shown for comparison.

| Method | Clip | | Flip | | Shift | |
|---|---|---|---|---|---|---|
| | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC |
| Jaccard | 0.843 ± 0.029 | 0.903 ± 0.027 | 0.843 ± 0.029 | 0.903 ± 0.027 | 0.843 ± 0.029 | 0.903 ± 0.027 |
| Lin | 0.832 ± 0.042 | 0.899 ± 0.025 | 0.832 ± 0.042 | 0.899 ± 0.024 | 0.832 ± 0.042 | 0.899 ± 0.024 |
| Resnik | 0.832 ± 0.043 | 0.900 ± 0.025 | 0.832 ± 0.043 | 0.900 ± 0.025 | 0.837 ± 0.040 | 0.899 ± 0.025 |

**Table 4. Model performance of semantic similarity measures across levels of simulated phenotypic sparsity and noise (see Fig 2).** Abbreviations: ACC–accuracy; AU-ROC–area under the receiver operator characteristics curve.

| Method | Noise 0.25 | | Noise 0.5 | | Noise 1 | | Noise 2 | | Noise 4 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC |
| Jaccard | 0.789 ± 0.033 | 0.855 ± 0.045 | 0.829 ± 0.048 | 0.895 ± 0.040 | 0.843 ± 0.029 | 0.903 ± 0.027 | 0.821 ± 0.012 | 0.896 ± 0.024 | 0.795 ± 0.056 | 0.875 ± 0.053 |
| Lin | 0.741 ± 0.059 | 0.830 ± 0.061 | 0.827 ± 0.040 | 0.889 ± 0.033 | 0.832 ± 0.042 | 0.899 ± 0.025 | 0.821 ± 0.028 | 0.896 ± 0.023 | 0.816 ± 0.032 | 0.885 ± 0.021 |
| Resnik | 0.739 ± 0.054 | 0.829 ± 0.063 | 0.832 ± 0.041 | 0.891 ± 0.030 | 0.837 ± 0.040 | 0.899 ± 0.025 | 0.824 ± 0.024 | 0.896 ± 0.025 | 0.805 ± 0.020 | 0.884 ± 0.023 |

https://doi.org/10.1371/journal.pcbi.1010959.t004

## Localized multi-task multi-kernel learning with hierarchical decomposition offers biological insight

Previously, we trained the model under the assumption that each kernel matrix (modality) contributes equally by assigning uniform global weights. The global weights for the best convex combination of kernel matrices can also be conveniently learnt from the data with wrapper methods, which have previously been demonstrated as promising tools to work with multi-source genomic data [35]. Here, we observed no significant difference in model performance between uniform weights and learnt global weights (Table 5). This is not surprising, as learning weights is most useful when we are working with a large set of kernel matrices and aim to achieve a sparse solution. If we have the prior knowledge to select a small set of adequate kernel matrices, uniform weights are usually sufficient. For theoretical guarantees, see Bach et al. [36].

Global weights assign the same weight to a kernel for all pairs of samples. Therefore, localized weights for each task in the input space may be more appropriate for precision medicine approaches [37]. We implemented hierarchical decomposition multi-task multi-kernel learning, which makes efficient use of our prior knowledge on channel taxonomy to learn localized weights for each task and each subset of tasks that are leaves of subtrees rooted at each node [38]. Thus, we could identify tasks and subsets of related tasks where the weight of our phenotypic similarity kernel was high. These tasks were likely to contain phenotypic information that contributed towards classification between GOF and LOF, which may be useful for downstream analysis. A dendrogram of voltage-gated sodium channels and their local weights is shown in S3 Fig. The model correctly identified channels with known or shared genotype-phenotype correlations, e.g. *SCN1A*, *SCN2A*, *SCN3A*, and *SCN5A*. A weighted linear combination of these subset kernel matrices was then used as a composite kernel matrix for the model, where the weights were learnt with wrapper methods. Here, the weight vector favoured the kernel matrix of the subset of all tasks (root, wt = 0.93). Thus, the model recognized that voltage-gated sodium channels were similar enough to be learnt together. Overall, hierarchical decomposition offered a modest improvement in model performance (Table 5).

## Finding difficult-to-predict variants in the reproducing kernel Hilbert space

Kernel-SVMs are considered 'black box' machine learning models. However, we may still be able to indirectly gather insight on model decisions. Fig 3A shows the histogram of projections for our model, as introduced by Cherkassky et al. [39]. Training data were projected onto the

**Table 5. Nested cross-validation (outer folds) model performance estimates for each multi-kernel learning method.** Uniform weights correspond to the sum of kernel matrices. For global weights, a wrapper method is used to learn the optimal weighted sum of kernel matrices. Hierarchical decomposition learns both localized weights and an optimal weighted sum of subset kernel matrices. The highest mean metric for each method is shown in bold.

| Method | ACC | $F_1$ score | Cohen's κ | MCC | AU-PRC | AU-ROC |
|---|---|---|---|---|---|---|
| Uniform weights | 0.837 ± 0.035 | 0.824 ± 0.029 | 0.673 ± 0.067 | 0.677 ± 0.064 | 0.837 ± 0.039 | 0.890 ± 0.023 |
| Global weights | 0.843 ± 0.038 | 0.832 ± 0.029 | 0.684 ± 0.071 | 0.691 ± 0.069 | **0.880 ± 0.046** | **0.915 ± 0.018** |
| Hierarchical decomposition | **0.853 ± 0.016** | **0.838 ± 0.028** | **0.705 ± 0.031** | **0.715 ± 0.018** | 0.858 ± 0.058 | 0.912 ± 0.025 |

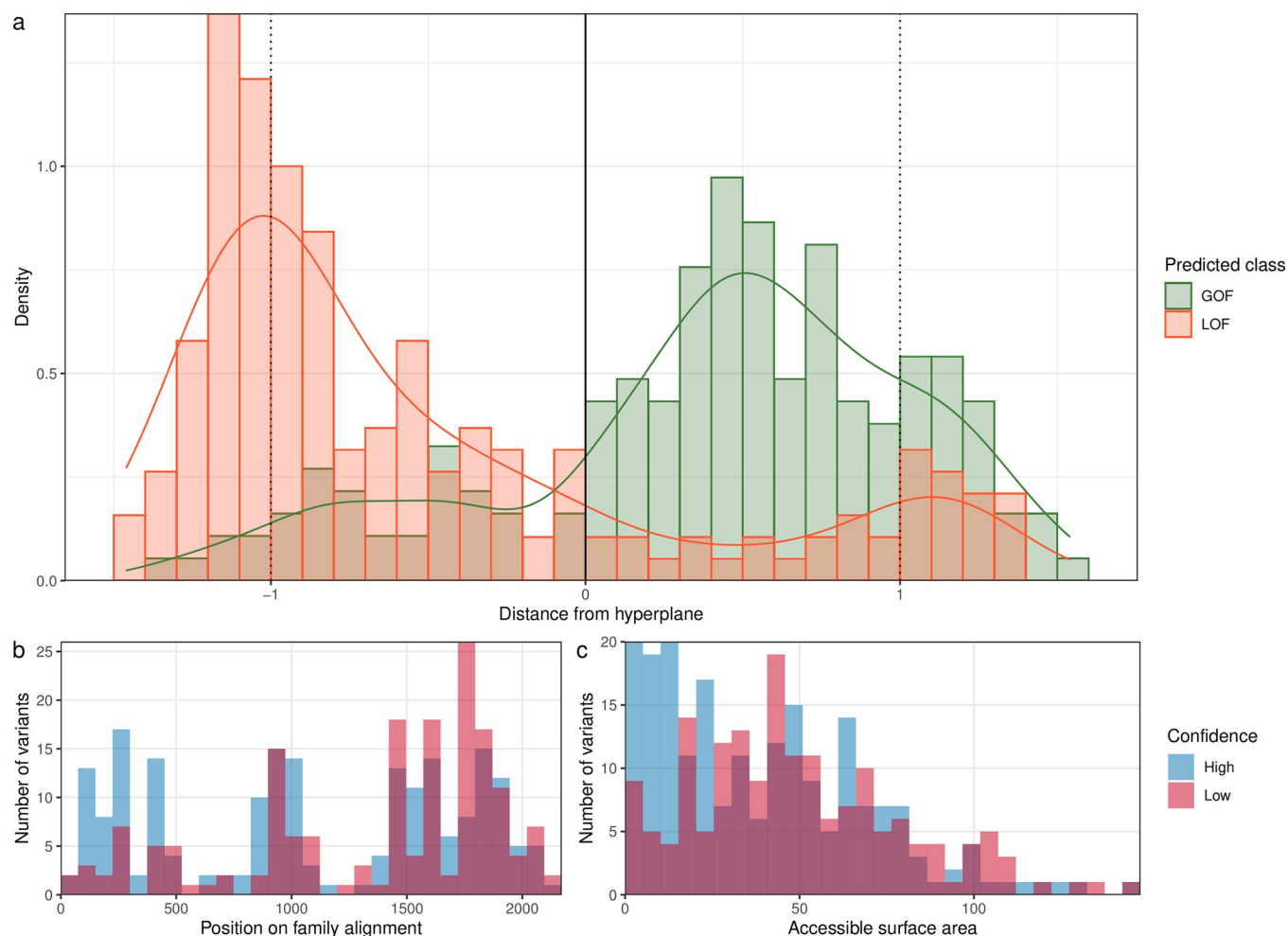https://doi.org/10.1371/journal.pcbi.1010959.t005

**Fig 3.** *a*: Cherkassky histogram of projections for the phenotypic multi-task multi-kernel learning model. *b-c*: Features that were significantly different between variants predicted with low confidence (distance to hyperplane $\leq 0.5$) and high confidence (distance to hyperplane $\geq 0.5$). The accessible surface area feature is based on surrogate predictions by NetSurfP-2.0 and is given in Å [40].

normal vector and plotted by their feature space distance to the optimally separating hyperplane. Samples within the range $\{x \in \mathbb{R} : -1 \leq x \leq 1\}$ were support vectors. If a sample is close to the hyperplane, model prediction confidence decreases. The set of samples within an arbitrary distance of $\{x \in \mathbb{R} : -0.5 \leq x \leq 0.5\}$ ("low confidence") was compared to the set of samples outside this range ("high confidence") by feature correlation with an unpaired two-samples Wilcoxon test. We adjusted p-values for multiple testing with the Benjamini-Hochberg method. Features that were significantly ($\alpha = 0.05$) associated with variants that were hard to predict ("low confidence") are shown in Fig 3B and 3C. This included variants in hotspots along the channel family sequence alignment corresponding to repeat domain IV, as well as variants with a high accessible surface area. For the final model, "low confidence" variants (n = 171) were predicted with a mean accuracy of 0.760, while "high confidence" variants (n = 204) were predicted with a mean accuracy of 0.931.

## Discussion

Learning with phenotypic similarity as features represents a natural extension of previous sequence- and structure-based machine learning algorithms [27,28]. Previous studies have

demonstrated that clinical features represent an important source of information that may be sufficient to clearly delineate between GOF and LOF in several channelopathies [41,42]. Here, we introduce the concept of extending semantic similarity analysis of human phenotype ontology (HPO) terms towards kernel-based machine learning methods. Beyond the gain in predictive performance, this method has several distinct advantages: i) HPO terms have become the standardized language in clinical genetics. They have been used to structure phenotypic information in large-scale sequencing consortia [34], enable deep phenotyping at the point of care [43], and drive phenotype matching software such as Exomiser [44]; ii) Conventionally, all features used for model training must be available at time of prediction. That is, presence or absence of all clinical information would have to be explicitly coded which quickly becomes intractable for complex models. The underlying knowledge graph of the ontology and use of similarity as features solves this issue; iii) Likewise, genotype-phenotype correlations do not have to be explicitly provided to the model but are instead learnt from the data. Identifying channels with high phenotypic similarity matrix weight offers insight for downstream analysis.

Extending our previous multi-task learning framework with multiple kernels allows us to include more information in the feature space. As the kernel matrix represents the information bottleneck for the analysis method [45], including more information has recently contributed towards data integration in multi-omics approaches [46,47]. Accordingly, we observed a large gain in performance for our model when compared to conventional baseline and current state-of-the-art methods. Model performance was robust across different semantic similarity measures. For the final model, we chose the Jaccard index as it represents a computationally efficient method that is guaranteed to result in a similarity matrix that fulfils model assumptions, avoiding theoretical consistency issues that may arise with the Lin and Resnik measures [31]. Encouragingly, model performance remained largely stable across simulated sparse and noisy phenotypes which more closely reflects the reality of routinely collected health data such as electronic health records and heterogeneous patient cohorts.

Localized multi-kernel learning by hierarchical decomposition offered insight into implicit phenotype correlations, identifying channels where clinical information is of additional use. The method is also capable of refining prior knowledge on task similarity with latent task similarity learnt from the data. Model performance only increased modestly when compared to the global model as the tasks were overall similar to each other. While this is true for the family of voltage-gated sodium channels [48,49], this general framework is a promising approach that we expect to scale well to larger task structures (such as unified models of different ion channel families) and to taxonomies with unequal edge lengths [38]. Lastly, feature correlation from the histogram of projections of our model identified two features that are associated with hard-to-predict variants: Location in repeat domain IV and a high residue accessible surface area. This information reflects our current mechanistic and structural understanding of sodium channels [48], and helps users identify whether their prediction is in a low-confidence region. That said, localized multi-kernel learning failed to accurately identify some channels with known genotype-phenotype correlations. For *SCN4A*, the clinical spectrum of myopathies is well known [1,50]. For *SCN8A*, recent work by Johannesen et al. has elegantly demonstrated that age at seizure onset and type of epilepsy can delineate between functional effects [24]. It is likely that our model currently does not have the resolution necessary to accurately discriminate based on detailed concepts such as myopathy subtypes or age at seizure onset, as the training data set contained only coarse key phenotypes. Additionally, *SCN9A* was learnt with low phenotypic weight likely due to class imbalance as the training data only included a single *SCN9A*-LOF. Phenotypic learning will benefit from large deeply phenotyped cohorts and more granular representation of concepts within the ontology [34].

While our model represents a large step forward in functional variant prediction, several limitations still apply. Our model predicts functional effect at the alpha subunit level, but does

not include data from auxiliary beta subunits (*SCN1B-SCN3B*) or other modulator proteins (e.g. *FGF12*) which may themselves affect channel expression levels, gating kinetics, or neuronal network activity [51–53]. The model is trained on data from whole-cell patch-clamp recordings of mammalian cells but includes no further information on experimental metadata. Hence, results may differ depending on the choice of expression system [14]. Regarding the results of internal and external validation, filtering variants with a clear net gain- or loss-of-function introduces selection bias by neglecting those variants with complex or mixed functional effects. For these variants in particular, electrophysiology remains the gold standard for evaluating functional effects. Most importantly, in silico tools like our model must not be considered strong or causal evidence by the genetic testing standards of the American College of Medical Genetics and Genomics [54], and should not be solely used to establish variant pathogenicity or to directly inform diagnosis and management of affected individuals. These novel tools require further validation and should currently be considered as prediction engines for research use, not as treatment decision support systems.

Learning with phenotypic similarity represents a novel step from harmonized ontological representations of clinical concepts towards kernel-based machine learning methods. Our framework is robust to phenotypic noise, computationally efficient, scalable to different channelopathies, and outperforms other state-of-the-art methods. The model generates supporting evidence for variant interpretation, and may aid clinicians and biophysical researchers in generating new clinical and experimental hypotheses.

## Materials and methods

### Data

Recently, Brunklaus et al. conducted a systematic literature search of variants in voltage-gated sodium channels that were functionally assessed in whole-cell patch-clamp recordings of mammalian cells [29]. We refer to their Supplementary Methods for the criteria used to assign functional labels to each variant. Of 437 variants, 62 variants were similar to wildtype function or did not have a clear overall gain- or loss-of-function. We removed these variants from further analysis. Our training data set included 375 unique non-synonymous missense variants across the 9 voltage-gated sodium channels ($Na_V1.1$-$Na_V1.9$). 164 variants were labelled as GOF, 211 variants were labelled as LOF.

### Features

Each variant was identified by its channel, sequence position, and amino acid substitution. Sequence- and structure-based features were extracted and pre-processed as previously described [28]. These 62 features included amino acid physicochemical properties, radicality of substitution, paralog and ortholog conservation, secondary and tertiary structure, residue accessibility, disordered regions, and protein-protein interaction or binding sites (S2 Table). Variant-level information on primary disease and reference for phenotypes were extracted from the training data set and mapped onto Phenotype MIM numbers of the Online Mendelian Inheritance in Man (OMIM) [55]. Each OMIM ID is associated with a set of human phenotype ontology (HPO) terms which represent a standardized vocabulary of phenotypic features in human disease [44]. These mappings were requested via the HPO REST API (last accessed 22/June/2022 13:35 pm, release v2022-06-11) and resulted in an average of 9 HPO terms per variant (median 9, range 3–42 terms). All parents of the HPO terms in the directed acyclic graph from the parent ← child relationships were added to the respective HPO term sets, which resulted in 446 unique terms across the set of all variants, with each variant enriched to an average of 40 terms (median 38, range 12–226). Ontology analysis was performed with the ontologyX suite [56].

We considered three different measures for semantic similarity analysis. For the Jaccard index, we take the ratio of the number of elements in the intersection and the number of elements in the union of the two variant sets. A more rigorous definition and proof of positive definiteness of the function has been provided by Gärtner et al. [30]. For the Resnik measure, the similarity between two variants is given as the information content (IC) of their most informative common ancestor (MICA) term [32]. The IC is defined as $-\log_2(f)$ where $f$ is the term frequency over all sets and the MICA is the common ancestor term with the highest IC. Lastly, the Lin measure extends the Resnik measure by incorporating the IC of the individual terms [57]. Further detail and a visual overview on phenotypic feature extraction and processing are shown in S1 Fig.

Both Resnik and Lin measures do not lead to positive semi-definite (PSD) similarity matrices in ontologies with multiple ancestors such as the HPO, where terms may have more than one parent term each. Before proceeding with kernel-based learning, we approximated the nearest PSD matrix with three methods suggested by Chen et al. [31]. In short, these are: i) *clip*, where each negative eigenvalue in our similarity matrix S is set to zero; ii) *flip*, where we flip the sign of each negative eigenvalue, which is equivalent to replacing the original eigenvalues of S with its singular values; iii) and *shift*, where we shift the spectrum of S by the absolute value of the minimum eigenvalue. For more rigorous definitions and an in-depth comparison of each method, see Chen et al. [31].

## Models

When we consider phenotypic semantic similarity analysis as a kernel function and assert its positive semidefinite properties, we satisfy Mercer's condition and can make implicit use of the reproducing kernel Hilbert space (RKHS) induced by our so-called phenotype kernel. We can avoid explicit representation of each variant in a Euclidean space and instead consider only their inner product, also thought of as pairwise similarity in the high-dimensional phenotypic feature space. This similarity-based learning represents a natural extension of semantic similarity analysis towards kernel-based machine learning methods.

For each variant, the vector of sequence- and structure-based features is transformed into pairwise similarities with a radial basis function (RBF) kernel $K_n$ with $k\left(x_i, x_j\right) = exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2}\right) \forall x_i, x_j \in X$ [58]. Task similarity, i.e. the hierarchical taxonomy of voltage-gated sodium channels across the channel superfamily, is represented as an alignment-based tree distance measure and enables multi-task (parallel transfer) learning. Both methods have been previously described [28]. We can consider each of these three data input types of our variants (channel, sequence/structure, and phenotype) as different measurement modalities with individual kernel matrices for each modality. First, each kernel matrix is normalized in feature space:

$$\hat{k}(x, z) = \langle \hat{\phi}(x), \hat{\phi}(z) \rangle = \frac{k(x, z)}{\sqrt{k(x, x)k(z, z)}}$$

Multiple kernel learning then allows us to represent these modalities in a single kernel matrix $K$, either by assigning uniform weights or by learning a linear combination of matrices $K_m$ with weights $\beta_m$ such that:

$$K = \sum_{m=1}^{M} \beta_m K_m, \ \beta \geq 0.$$

Global weights assign the same weight over the whole input space, which may not be desirable. Several methods of localized multiple kernel learning have instead been proposed [37].

We implemented hierarchical decomposition multi-task multi-kernel learning (MTMKL) by Widmer et al. [38]. This method makes use of prior knowledge in the form of a tree structure $G$, which in our setting is the taxonomy that relates voltage-gated sodium channels, where each channel is a task $t_i$ corresponding to a leaf in $G$. Given the set of tasks $T = \{t_1, t_2, \ldots, t_T\}$, let the subsets of tasks $S_i \subseteq T$ be the leaves of the sub-tree rooted at each node $n$ of $G$, that is *leaves* $(n) = \{l | l$ is descendant of $n\}$ [38]. The following definition for the kernel function is given:

$$\hat{K} = \sum_{n_i \in G} \beta_i K_{leaves(n_i)}$$

A visualization of these kernel matrices is shown in S2 Fig. This method also offers additional insight [38]. First, the localized kernel weights of each task allow us to see which modality is important for which task. Secondly, we can refine our prior task similarity by learning latent task similarity from the data. If two tasks $t_k$ and $t_l$ are present in subsets $S_i$ that are assigned weights $\beta_i$, we can define their latent similarity by summing up the weights of their shared task sets where the set of tasks sets containing task $t_k$ is defined as $T_k = \{S | t_k \in S \land S \in T\}$, where $T$ is the set of all tasks, such that the pairwise similarity of tasks $\gamma_{k,l}$ is given as:

$$\gamma_{k,l} = \sum_{S_i \in T_k \cap T_l} \beta_i$$

We trained a support-vector machine (C-SVM) on the resulting kernel matrix. Hyperparameter tuning was performed via grid search across cost C = {1e−4, 1e−2,. . .,1e4}, the RBF kernel hyperparameter σ = {1e−5, 1e−4,. . .,1e0}, and hierarchical multi-task learning baseline similarity parameter $a$ = {1, 3, 5, 10, 100}. The choice of semantic similarity measure and PSD matrix approximation were included in the set of hyperparameters. Hyperparameter tuning and model assessment were done with nested $k$-fold cross validation, $k$ = 5, where the model is trained and tested on different random sub-partitions of the data throughout multiple exhaustive iterations to obtain an estimate of the generalization performance. For illustrative purposes, a pseudocode and graphical representation of $k$-fold cross validation is provided in our previous work [28]. We report the following metrics: i) Accuracy, the number of correct classifications divided by the number of all classifications; ii) Sensitivity, or true positive rate (TPR); iii) Specificity, or true negative rate (TNR); iv) Matthews correlation coefficient (MCC), a symmetric correlation coefficient between true and predicted classes; iii) Cohen's Kappa ($\kappa$), a measure of agreement between true and predicted classes; iv) $F_1$ score, the harmonic mean of precision and recall; v) The area under the receiver operator characteristics (AU-ROC) curve; vi) and the area under the precision-recall curve (AU-PRC).

Further model information is available in our AIMe Registry report (aime.report/xmSKHj).

## Validation

For baseline comparison, we trained one SVM on variants from each task (channel-specific model, Dirac) and one SVM on variants from all tasks (global model, Union). For internal validation, we implemented gradient boosting as described by Heyne et al. [27], training and testing their method on the same splits of our data set. We also implemented the decision rule model proposed by Brunklaus et al., assigning labels based on missense pairs of identical paralog variants across the sodium channel family [29]. For external validation, we conducted a systematic literature search of PubMed with the search string (((SCN1A) OR (SCN2A) OR (SCN3A) OR (SCN4A) OR (SCN5A) OR (SCN8A) OR (SCN9A) OR (SCN10A) OR (SCN11A)) AND ((gain-of-function) OR (loss-of-function)) OR ((GOF) OR (LOF)) AND

((electrophysiology) OR (patch clamp) OR (voltage clamp))) AND (("2021/04/28"[Date—Publication]: "2022/07/13"[Date—Publication])). This query captures all publications after the last update of the training data set. 55 studies were identified and screened for missense variants in voltage-gated sodium channels with experimental evidence of whole-cell patch-clamp recordings that yielded an overall gain- or loss-of-function effect (n = 91). Variants that were present in the training data set, i.e. those known from previous functional experiments, were removed from further analysis (n = 61). Predictions were obtained from our own model and the web interfaces of Heyne et al. (funNCion, http://funncion.broadinstitute.org/), and Brunklaus et al. (SCN viewer, https://scn-viewer.broadinstitute.org/).

## Software

This study was carried out in the R programming language, version 4.1.0, with RStudio, version 1.4.1106.

## Supporting information

**S1 Fig. Human phenotype ontology (HPO) terms and phenotypic learning.** *a*: Clinical descriptions from medical records or literature are harmonized by extracting relevant concepts and mapping them to the standardized vocabulary of HPO terms [1]. *b*: Each term represents a node in a directed acyclic graph. Edges denote directional IS-A (parent-child) relationships. During propagation, for each term in the set of terms, each parent of a term is included in the set of terms until the root node (HP:0000001 All) is included. Thus, the underlying knowledge graph allows the model to understand that, e.g., HP:0001249 Intellectual disability IS-A HP:0012759 Neurodevelopmental abnormality. Nodes are colored by term, with gray nodes not being present in the set of terms. *c*: Given two sets of terms, their pairwise semantic similarity is calculated as described above (Methods). In this example, the most informative common ancestor (MICA) method is shown. The information content (IC) of this node is $-\log_2(f)$ where $f$ is the frequency of the term over all sets of terms. In Resnik's measure, the IC of the MICA is the pairwise similarity between the two sets of terms. *d*: The pairwise phenotypic similarity for each pair of n observations (variants) in the training data set is represented by a *n*-by-*n* square matrix. This matrix is a Gram matrix (or kernel matrix) if it is positive semi-definite. Here, a simplified low-dimensional representation of the resulting phenotypic similarity feature space is shown. The bold line is the hyperplane of our support-vector machine classifier, with the dashed lines being the soft margin.
(PDF)

**S2 Fig. Kernel matrix visualization as correlation plots ordered by hierarchical clustering.** *a*: Sequence- and structure-based pairwise similarity matrix. *b*: Task similarity matrix. *c*: Multi-task learning kernel matrix. *d*: Phenotypic similarity matrix. *e*: Multi-task multi-kernel learning matrix.
(PDF)

**S3 Fig. Dendrogram of hierarchical decomposition multi-task multi-kernel learning for voltage-gated sodium channels.** Node size and weight labels on the branches correspond to localized phenotypic similarity kernel weights. Weights were assumed to be uniform for *SCN11A*, as the sample size of the leaf subset was too small. Leaf-to-leaf distance on the dendrogram corresponds to the task similarity used for multi-task learning, with similar tasks clustered together.
(PDF)

**S1 Table. Variants used for external validation.** Details on data curation and validation procedure are available in the Methods section. Primary disease corresponds to the core phenotype, while phenotype terms refers to the information provided to the MKL model. For the decision rule, exact matches are shown in bold. Abbreviations: BrS–Brugada syndrome; DEE–developmental and epileptic encephalopathy; DS–Dravet syndrome; ERS–Early repolarization syndrome; ESES—electrical status epilepticus during slow-wave sleep; GOF–gain-of-function; LOF–loss-of-function; LQT3 –long-QT syndrome 3; PMC–paramyotonia congenita; SeLIE–self-limited (familial) infantile *epilepsy*.
(PDF)

**S2 Table. List and description of sequence- and structure-based features, sorted by column vectors in the training data set.** Raw and preprocessed feature tables are available as part of our 'Availability of data and materials' statement.
(PDF)

## Author Contributions

**Formal analysis:** Christian Malte Boßelmann, Nico Pfeifer.

**Methodology:** Christian Malte Boßelmann.

**Resources:** Holger Lerche, Nico Pfeifer.

**Software:** Christian Malte Boßelmann.

**Supervision:** Holger Lerche, Nico Pfeifer.

**Validation:** Ulrike B. S. Hedrich.

**Visualization:** Christian Malte Boßelmann.

**Writing – original draft:** Christian Malte Boßelmann.

**Writing – review & editing:** Ulrike B. S. Hedrich, Holger Lerche, Nico Pfeifer.

## References

1. Lerche H, Mitrovic N, Jurkat-Rott K, Lehmann-Horn F. Chapter 8 Sodium channelopathies in skeletal muscle and brain. In: Reisin RC, Nuwer MR, Hallett M, Medina C, editors. Supplements to Clinical Neurophysiology. 54: Elsevier; 2002. p. 62–9.

2. Symonds JD, Zuberi SM, Stewart K, McLellan A, O'Regan M, MacLeod S, et al. Incidence and phenotypes of childhood-onset genetic epilepsies: a prospective population-based national cohort. Brain. 2019; 142(8):2303–18. https://doi.org/10.1093/brain/awz195 PMID: 31302675; PubMed Central PMCID: PMC6658850.

3. Symonds JD, Elliott KS, Shetty J, Armstrong M, Brunklaus A, Cutcutache I, et al. Early childhood epilepsies: epidemiology, classification, aetiology, and socio-economic determinants. Brain. 2021; 144 (9):2879–91. https://doi.org/10.1093/brain/awab162 PMID: 34687210; PubMed Central PMCID: PMC8557326.

4. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). Heart Rhythm. 2011; 8(8):1308–39. https://doi.org/10.1016/j.hrthm.2011.05.020 PMID: 21787999.

5. Fertleman CR, Ferrie CD, Aicardi J, Bednarek NA, Eeg-Olofsson O, Elmslie FV, et al. Paroxysmal extreme pain disorder (previously familial rectal pain syndrome). Neurology. 2007; 69(6):586–95. https://doi.org/10.1212/01.wnl.0000268065.16865.5f PMID: 17679678.

6. Kellenberger S, West JW, Catterall WA, Scheuer T. Molecular analysis of potential hinge residues in the inactivation gate of brain type IIA Na+ channels. J Gen Physiol. 1997; 109(5):607–17. https://doi.org/10.1085/jgp.109.5.607 PMID: 9154907; PubMed Central PMCID: PMC2217067.

7. Kellenberger S, West JW, Scheuer T, Catterall WA. Molecular analysis of the putative inactivation particle in the inactivation gate of brain type IIA Na+ channels. J Gen Physiol. 1997; 109(5):589–605. https://doi.org/10.1085/jgp.109.5.589 PMID: 9154906; PubMed Central PMCID: PMC2217064.

8. McPhee JC, Ragsdale DS, Scheuer T, Catterall WA. A critical role for transmembrane segment IVS6 of the sodium channel alpha subunit in fast inactivation. J Biol Chem. 1995; 270(20):12025–34. https://doi.org/10.1074/jbc.270.20.12025 PMID: 7744852.

9. Jurkat-Rott K, Groome J, Lehmann-Horn F. Pathophysiological role of omega pore current in channelopathies. Front Pharmacol. 2012; 3:112. Epub 20120611. https://doi.org/10.3389/fphar.2012.00112 PMID: 22701429; PubMed Central PMCID: PMC3372090.

10. Lehmann-Horn F, Jurkat-Rott K. Voltage-gated ion channels and hereditary disease. Physiol Rev. 1999; 79(4):1317–72. https://doi.org/10.1152/physrev.1999.79.4.1317 PMID: 10508236.

11. Lauxmann S, Boutry-Kryza N, Rivier C, Mueller S, Hedrich UB, Maljevic S, et al. An SCN2A mutation in a family with infantile seizures from Madagascar reveals an increased subthreshold Na(+) current. Epilepsia. 2013; 54(9):e117–21. Epub 20130612. https://doi.org/10.1111/epi.12241 PMID: 23758435.

12. Mantegazza M, Cestele S, Catterall WA. Sodium channelopathies of skeletal muscle and brain. Physiol Rev. 2021; 101(4):1633–89. Epub 20210326. https://doi.org/10.1152/physrev.00025.2020 PMID: 33769100; PubMed Central PMCID: PMC8989381.

13. Hubner CA, Jentsch TJ. Ion channel diseases. Hum Mol Genet. 2002; 11(20):2435–45. https://doi.org/10.1093/hmg/11.20.2435 PMID: 12351579.

14. Makinson CD, Dutt K, Lin F, Papale LA, Shankar A, Barela AJ, et al. An Scn1a epilepsy mutation in Scn8a alters seizure susceptibility and behavior. Exp Neurol. 2016; 275 Pt 1:46–58. Epub 20150926. https://doi.org/10.1016/j.expneurol.2015.09.008 PMID: 26410685; PubMed Central PMCID: PMC4688066.

15. de Kovel CG, Meisler MH, Brilstra EH, van Berkestijn FM, van 't Slot R, van Lieshout S, et al. Characterization of a de novo SCN8A mutation in a patient with epileptic encephalopathy. Epilepsy Res. 2014; 108(9):1511–8. Epub 20140904. https://doi.org/10.1016/j.eplepsyres.2014.08.020 PMID: 25239001; PubMed Central PMCID: PMC4490185.

16. Dichgans M, Freilinger T, Eckstein G, Babini E, Lorenz-Depiereux B, Biskup S, et al. Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. Lancet. 2005; 366 (9483):371–7. https://doi.org/10.1016/S0140-6736(05)66786-4 PMID: 16054936.

17. Brunklaus A, Brunger T, Feng T, Fons C, Lehikoinen A, Panagiotakaki E, et al. The gain of function SCN1A disorder spectrum: novel epilepsy phenotypes and therapeutic implications. Brain. 2022. Epub 20220613. https://doi.org/10.1093/brain/awac210 PMID: 35696452.

18. Deschenes I, Baroudi G, Berthet M, Barde I, Chalvidan T, Denjoy I, et al. Electrophysiological characterization of SCN5A mutations causing long QT (E1784K) and Brugada (R1512W and R1432G) syndromes. Cardiovasc Res. 2000; 46(1):55–65. https://doi.org/10.1016/s0008-6363(00)00006-7 PMID: 10727653.

19. Chen Q, Kirsch GE, Zhang D, Brugada R, Brugada J, Brugada P, et al. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. Nature. 1998; 392(6673):293–6. https://doi.org/10.1038/32675 PMID: 9521325.

20. Wang Q, Shen J, Splawski I, Atkinson D, Li Z, Robinson JL, et al. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. Cell. 1995; 80(5):805–11. https://doi.org/10.1016/0092-8674(95)90359-3 PMID: 7889574.

21. Bennett PB, Yazawa K, Makita N, George AL Jr. Molecular mechanism for an inherited cardiac arrhythmia. Nature. 1995; 376(6542):683–5. https://doi.org/10.1038/376683a0 PMID: 7651517.

22. Cox JJ, Reimann F, Nicholas AK, Thornton G, Roberts E, Springell K, et al. An SCN9A channelopathy causes congenital inability to experience pain. Nature. 2006; 444(7121):894–8. https://doi.org/10.1038/nature05413 PMID: 17167479; PubMed Central PMCID: PMC7212082.

23. Wolff M, Johannesen KM, Hedrich UBS, Masnada S, Rubboli G, Gardella E, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. Brain. 2017; 140 (5):1316–36. https://doi.org/10.1093/brain/awx054 PMID: 28379373.

24. Johannesen KM, Liu Y, Koko M, Gjerulfsen CE, Sonnenberg L, Schubert J, et al. Genotype-phenotype correlations in SCN8A-related disorders reveal prognostic and therapeutic implications. Brain. 2021. Epub 20210825. https://doi.org/10.1093/brain/awab321 PMID: 34431999.

25. Knowles JK, Helbig I, Metcalf CS, Lubbers LS, Isom LL, Demarest S, et al. Precision medicine for genetic epilepsy on the horizon: Recent advances, present challenges, and suggestions for continued progress. Epilepsia. 2022. Epub 20220618. https://doi.org/10.1111/epi.17332 PMID: 35716052.

26. Cooper GM, Shendure J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. Nat Rev Genet. 2011; 12(9):628–40. Epub 20110818. https://doi.org/10.1038/nrg3046 PMID: 21850043.

27. Heyne HO, Baez-Nieto D, Iqbal S, Palmer DS, Brunklaus A, May P, et al. Predicting functional effects of missense variants in voltage-gated sodium and calcium channels. Sci Transl Med. 2020; 12(556). https://doi.org/10.1126/scitranslmed.aay6848 PMID: 32801145.

28. Bosselmann CM, Hedrich UBS, Muller P, Sonnenberg L, Parthasarathy S, Helbig I, et al. Predicting the functional effects of voltage-gated potassium channel missense variants with multi-task learning. EBio-Medicine. 2022; 81:104115. Epub 20220624. https://doi.org/10.1016/j.ebiom.2022.104115 PMID: 35759918; PubMed Central PMCID: PMC9250003.

29. Brunklaus A, Feng T, Brunger T, Perez-Palma E, Heyne H, Matthews E, et al. Gene variant effects across sodium channelopathies predict function and guide precision therapy. Brain. 2022. Epub 20220117. https://doi.org/10.1093/brain/awac006 PMID: 35037686.

30. Gärtner T, Le QV, Smola AJ. A short tour of kernel methods for graphs. Technical Report. 2006.

31. Chen Y, Garcia EK, Gupta MR, Rahimi A, Cazzanti L. Similarity-based classification: Concepts and algorithms. Journal of Machine Learning Research. 2009; 10.

32. Resnik P. Using information content to evaluate semantic similarity in a taxonomy. Proceedings of the 14th international joint conference on Artificial intelligence—Volume 1; Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc.; 1995. p. 448–53.

33. McDavid A, Crane PK, Newton KM, Crosslin DR, McCormick W, Weston N, et al. Enhancing the power of genetic association studies through the use of silver standard cases derived from electronic medical records. PLoS One. 2013; 8(6):e63481. Epub 20130610. https://doi.org/10.1371/journal.pone.0063481 PMID: 23762230; PubMed Central PMCID: PMC3677889.

34. Lewis-Smith D, Parthasarathy S, Xian J, Kaufman MC, Ganesan S, Galer PD, et al. Computational analysis of neurodevelopmental phenotypes: Harmonization empowers clinical discovery. Hum Mutat. 2022. Epub 20220423. https://doi.org/10.1002/humu.24389 PMID: 35460582.

35. Wilson CM, Li K, Yu X, Kuan PF, Wang X. Multiple-kernel learning for genomic data mining and predic-tion. BMC Bioinformatics. 2019; 20(1):426. Epub 20190815. https://doi.org/10.1186/s12859-019-2992-1 PMID: 31416413; PubMed Central PMCID: PMC6694479.

36. Bach FR. Consistency of the Group Lasso and Multiple Kernel Learning. J Mach Learn Res. 2008; 9:1179–225.

37. Gönen M, Alpaydın E. Multiple Kernel Learning Algorithms. J Mach Learn Res. 2011; 12(null):2211–68.

38. Widmer C, Toussaint NC, Altun Y, Rätsch G. Inferring latent task structure for Multitask Learning by Multiple Kernel Learning. BMC Bioinformatics. 2010; 11(8):S5. https://doi.org/10.1186/1471-2105-11-S8-S5 PMID: 21034430

39. Cherkassky V, Dhar S. Simple Method for Interpretation of High-Dimensional Nonlinear SVM Classifica-tion Models 2010. 267–72 p.

40. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sonderby CK, et al. NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. Proteins. 2019; 87 (6):520–7. Epub 20190309. https://doi.org/10.1002/prot.25674 PMID: 30785653.

41. Berecki G, Howell KB, Heighway J, Olivier N, Rodda J, Overmars I, et al. Functional correlates of clinical phenotype and severity in recurrent SCN2A variants. Commun Biol. 2022; 5(1):515. Epub 20220530. https://doi.org/10.1038/s42003-022-03454-1 PMID: 35637276; PubMed Central PMCID: PMC9151917.

42. Crawford K, Xian J, Helbig KL, Galer PD, Parthasarathy S, Lewis-Smith D, et al. Computational analysis of 10,860 phenotypic annotations in individuals with SCN2A-related disorders. Genet Med. 2021; 23 (7):1263–72. Epub 20210317. https://doi.org/10.1038/s41436-021-01120-1 PMID: 33731876; PubMed Central PMCID: PMC8257493.

43. Havrilla JM, Singaravelu A, Driscoll DM, Minkovsky L, Helbig I, Medne L, et al. PheNominal: an EHR-integrated web application for structured deep phenotyping at the point of care. BMC Med Inform Decis Mak. 2022; 22(Suppl 2):198. Epub 20220728. https://doi.org/10.1186/s12911-022-01927-1 PMID: 35902925; PubMed Central PMCID: PMC9335954.

44. Kohler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, Vasilevsky NA, et al. The Human Phenotype Ontology in 2021. Nucleic Acids Res. 2021; 49(D1):D1207–17. https://doi.org/10.1093/nar/gkaa1043 PMID: 33264411; PubMed Central PMCID: PMC7778952.

45. Shawe-Taylor J, Cristianini N. Basic concepts. Kernel Methods for Pattern Analysis. Cambridge: Cam-bridge University Press; 2004. p. 1–2.

46. Huang S, Chaudhary K, Garmire LX. More Is Better: Recent Progress in Multi-Omics Data Integration Methods. Front Genet. 2017; 8:84. Epub 20170616. https://doi.org/10.3389/fgene.2017.00084 PMID: 28670325; PubMed Central PMCID: PMC5472696.

47. Speicher NK, Pfeifer N. Integrating different data types by regularized unsupervised multiple kernel learning with application to cancer subtype discovery. Bioinformatics. 2015; 31(12):i268–75. https://doi.org/10.1093/bioinformatics/btv244 PMID: 26072491; PubMed Central PMCID: PMC4765854.

48. Ahern CA, Payandeh J, Bosmans F, Chanda B. The hitchhiker's guide to the voltage-gated sodium channel galaxy. J Gen Physiol. 2016; 147(1):1–24. https://doi.org/10.1085/jgp.201511492 PMID: 26712848; PubMed Central PMCID: PMC4692491.

49. de Lera Ruiz M, Kraus RL. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. J Med Chem. 2015; 58(18):7093–118. Epub 20150514. https://doi.org/10.1021/jm501981g PMID: 25927480.

50. Jurkat-Rott K, Holzherr B, Fauler M, Lehmann-Horn F. Sodium channelopathies of skeletal muscle result from gain or loss of function. Pflugers Arch. 2010; 460(2):239–48. Epub 20100317. https://doi.org/10.1007/s00424-010-0814-4 PMID: 20237798; PubMed Central PMCID: PMC2883924.

51. Angsutararux P, Zhu W, Voelker TL, Silva JR. Molecular Pathology of Sodium Channel Beta-Subunit Variants. Front Pharmacol. 2021; 12:761275. Epub 20211119. https://doi.org/10.3389/fphar.2021.761275 PMID: 34867379; PubMed Central PMCID: PMC8640220.

52. Martinez-Moreno R, Selga E, Riuro H, Carreras D, Parnes M, Srinivasan C, et al. An SCN1B Variant Affects Both Cardiac-Type (NaV1.5) and Brain-Type (NaV1.1) Sodium Currents and Contributes to Complex Concomitant Brain and Cardiac Disorders. Front Cell Dev Biol. 2020; 8:528742. Epub 20200929. https://doi.org/10.3389/fcell.2020.528742 PMID: 33134290; PubMed Central PMCID: PMC7550680.

53. Seiffert S, Pendziwiat M, Bierhals T, Goel H, Schwarz N, van der Ven A, et al. Modulating effects of FGF12 variants on NaV1.2 and NaV1.6 being associated with developmental and epileptic encephalopathy and Autism spectrum disorder: A case series. EBioMedicine. 2022; 83:104234. Epub 20220824. https://doi.org/10.1016/j.ebiom.2022.104234 PMID: 36029553; PubMed Central PMCID: PMC9429545.

54. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015; 17(5):405–24. Epub 20150305. https://doi.org/10.1038/gim.2015.30 PMID: 25741868; PubMed Central PMCID: PMC4544753.

55. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. Nucleic Acids Res. 2015; 43(Database issue):D789–98. Epub 20141126. https://doi.org/10.1093/nar/gku1205 PMID: 25428349; PubMed Central PMCID: PMC4383985.

56. Greene D, Richardson S, Turro E. ontologyX: a suite of R packages for working with ontological data. Bioinformatics. 2017; 33(7):1104–6. https://doi.org/10.1093/bioinformatics/btw763 PMID: 28062448; PubMed Central PMCID: PMC5386138.

57. Lin D. An Information-Theoretic Definition of Similarity. Proceedings of the Fifteenth International Conference on Machine Learning: Morgan Kaufmann Publishers Inc.; 1998. p. 296–304.

58. Schölkopf B, Smola AJ. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond: The MIT Press; 2018.