Epilepsy & Behavior 56 (2016) 32–37

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Changing the approach to treatment choice in epilepsy using big data

Orrin Devinsky [a,*], Cynthia Dilley [b], Michal Ozery-Flato [c], Ranit Aharonov [c], Ya'ara Goldschmidt [c], Michal Rosen-Zvi [c], Chris Clark [b], Patty Fritz [b]

[a] Comprehensive Epilepsy Center, New York University Medical Center, 223 E. 34th Street, New York, NY 10016, USA
[b] UCB Pharma, 1950 Lake Park Dr., Smyrna, GA 30080, USA
[c] IBM Research, University of Haifa Campus, Mount Carmel, Haifa 3498825, Israel

ARTICLE INFO

Article history:
Received 16 October 2015
Revised 22 December 2015
Accepted 24 December 2015
Available online 29 January 2016

Keywords:
Antiepileptic drugs
Epilepsy
Treatment decision
Predictive model

ABSTRACT

Purpose: A UCB–IBM collaboration explored the application of machine learning to large claims databases to construct an algorithm for antiepileptic drug (AED) choice for individual patients.
Methods: Claims data were collected between January 2006 and September 2011 for patients with epilepsy >16 years of age. A subset of patient claims with a valid index date of AED treatment change (new, add, or switch) were used to train the AED prediction model by retrospectively evaluating an index date treatment for subsequent treatment change. Based on the trained model, a model-predicted AED regimen with the lowest likelihood of treatment change was assigned to each patient in the group of test claims, and outcomes were evaluated to test model validity.
Results: The model had 72% area under receiver operator characteristic curve, indicating good predictive power. Patients who were given the model-predicted AED regimen had significantly longer survival rates (time until a treatment change event) and lower expected health resource utilization on average than those who received another treatment. The actual prescribed AED regimen at the index date matched the model-predicted AED regimen in only 13% of cases; there were large discrepancies in the frequency of use of certain AEDs/combinations between model-predicted AED regimens and those actually prescribed.
Conclusions: Chances of treatment success were improved if patients received the model-predicted treatment. Using the model's prediction system may enable personalized, evidence-based epilepsy care, accelerating the match between patients and their ideal therapy, thereby delivering significantly better health outcomes for patients and providing health-care savings by applying resources more efficiently. Our goal will be to strengthen the predictive power of the model by integrating diverse data sets and potentially moving to prospective data collection.

1. Introduction

The clinician's ability to identify antiepileptic drug (AED) regimens that provide each patient with epilepsy with the best possible outcomes is a significant challenge. Clinical trials rarely provide the specificity for individual patient-centric decisions. Even existing epilepsy treatment guidelines provide recommendations on key aspects of care such as the treatment of new-onset epilepsy [1], treatment-resistant epilepsy [2], and epilepsy in patients with HIV/AIDS [3], but do not cater specifically to patients who vary, for example, by age, etiology, and socioeconomic status. Clinicians, therefore, often rely on trial and error.

While seizures in approximately 60% of patients respond to their first AED, another 15% spend 2–5 years finding an effective AED regimen; seizures in the remaining 25–30% are treatment-resistant [4,5]. This indicates a knowledge gap and significant unmet medical need regarding optimal AED choice for balancing symptom control and tolerability for individuals [6]. Indeed, the initial promise of personalized medicine was alluring in the management of epilepsy, but it has been slow to deliver [7,8].

With the increasing amount of clinical data and available AEDs, application of computer learning and data analysis may help physicians easily access the most relevant information to make treatment decisions. This approach takes mass quantities of structured and unstructured data from various sources and asks the computer to learn and return a set of structured answers based only on the most relevant data [9]. A clinical decision support (CDS) system for pediatric epilepsy was developed using a computer system to integrate expert opinion

Abbreviations: AED, antiepileptic drug; AUC, area under the curve; CDS, clinical decision support; ICD-9, International Classification of Diseases, Ninth Revision; ROC, receiver operator characteristic.

* Corresponding author. Tel.: +1 646 558 0801; fax: +1 646 385 7164.
E-mail addresses: od4@nyu.edu (O. Devinsky), Cynthia.Dilley@ucb.com (C. Dilley), OZERY@il.ibm.com (M. Ozery-Flato), ranitah1@gmail.com (R. Aharonov), YAARAG@il.ibm.com (Y. Goldschmidt), ROSEN@il.ibm.com (M. Rosen-Zvi), Chris.Clark@ucb.com (C. Clark), Patty.Fritz@ucb.com (P. Fritz).

http://dx.doi.org/10.1016/j.yebeh.2015.12.039
1525-5050/

with a configured knowledge base to make treatment recommendations [10,11]. A collaboration between UCB and IBM has been made to develop a predictive model offering personalized care for people with epilepsy that uses "cognitive computing" to analyze data involving thousands of longitudinal records, similar to a collaboration between IBM and Memorial Sloan Kettering Cancer Center developing a system to diagnose and treat an individual's cancer [9]. For epilepsy, this modeling approach leverages the increasing aggregation and collation of claims data to predict the chances of treatment success, defined by avoidance of hospitalization or treatment change, based on the similarity of the individual patient's characteristics to a larger patient population. Here, we present our initial findings in epilepsy.

2. Methods

2.1. Data source and patients

Medical, pharmacy, and hospital claims data were collected from all major regions of the United States between January 1, 2006 and September 31, 2011 from the IMS Health Surveillance Data Incorporated (SDI) medical claims database. This database was chosen because it broadly reflects the underlying population of patients with epilepsy including census-like geographic coverage, has full representation of third-party and government payers, and does not require continuous eligibility in a health plan, ensuring that varying socioeconomic status and patient movement across plans/payers each year does not impede the ability to track these patients over time. The SDI database provides de-identified patient data in compliance with Health Insurance Portability and Accountability Act regulations, thus making the study exempt from institutional review board review.

The SDI aggregates patient information from multiple provider sources, but might not capture all claims for an individual patient if providers that do not submit data to SDI were used. To control for this, we employed eligibility requirements for continuous reporting from the sources. Specifically, we required at least 80% continuous monthly eligibility (in 1-year windows) in any of the SDI pharmacy, physician, or hospital databases, and quarterly pharmacy eligibility for each patient. The analysis was performed on the longest eligible data period of the patient, requiring a minimum period of 2 years.

To capture data from patients with epilepsy rather than from patients receiving AEDs for other indications, we defined an analysis set of patients >16 years of age in January 2006, with at least one *International Classification of Diseases, Ninth Revision* (*ICD-9*) epilepsy diagnosis code (345.xx) or two seizure diagnosis codes (780.3x) at any time, as well as at least one claim for an AED from a pharmacy with 80% stability (existence of monthly pharmacy claims data) over the whole data period. The final inclusion criterion was that an index date (defined below) was identifiable for the patient.

2.2. Choice of index date

An index date was defined as the first date on which a treatment change occurred, where: (1) the new AED regimen switched one or more AEDs in the previous regimen or added an AED to the existing regimen (excluding the case of restarting a past treatment, and including the case of moving from no treatment to AED treatment); (2) the patient had an eligible period of ≥12 months before and after this date; (3) the patient had ≥3 months of pharmacy eligibility prior to this date; and (4) the treatment was unchanged during the 30-day period after the index date (to eliminate rescue medication in favor of chronic treatment).

The AED regimen following the treatment change event at the index date was defined as the index date AED regimen. The 12-month period before the index date was used to extract patient features, while the 12-month period after the index date was used to determine the outcome.

2.3. Choice of outcome

Because the primary symptom of epilepsy—seizures—is not captured in claims data, we used treatment change events as a proxy measure of seizure control and patient status. Intuitively, the need for treatment change indicates that the patient's current treatment regimen was suboptimal in terms of efficacy and/or tolerability. In order to be a valid index date, the treatment had to remain unchanged for at least 30 days post-index date. Therefore, an unsuccessful AED regimen was defined as any change other than a dose change (i.e., increase/decrease) or a complete withdrawal of any AED treatment in the subsequent 1–12 months after the change. Furthermore, longer-term stable treatment or a complete withdrawal from an AED therapy was assumed to indicate successful treatment.

2.4. Data analysis

The objective of the UCB–IBM collaboration was to retrospectively estimate the effectiveness of different treatment approaches using large observational data sets and use this information to predict successful treatments for individual patients. We used machine-learning methods to create a predictive algorithm estimating the success probability for a given patient and a specific treatment regimen. This algorithm was then used to predict the treatment regimen with the highest success probability for each patient (model-predicted AED regimen). The effects of using the algorithm's predictions were assessed in terms of treatment change rates and utilization of health-care services using an independent patient set.

Disjoint sets of 40,000 patients for training and 10,000 patients for testing were randomly selected from the patients meeting all study inclusion criteria. To increase the reliability of the results, we excluded patients whose index date treatment regimen was relatively rare (i.e., occurred <50 times in the training population).

2.4.1. Building and testing the model

We used the training set to train a predictive model which, given the patient data 12-month pre-index date and the index date AED regimen, predicts the probability of success. The predictive model was based on features extracted from the data, but not costs or health-care utilization features. From each patient's record, roughly 5000 features were extracted, which include patient features (related to the patient and recorded prior to the index date, e.g., demographics, medication history, comorbidities, *ICD-9* codes), treatment features (related to the index date AED regimen—the one to be evaluated—e.g., the number and type of distinct drugs at index date), and patient-treatment features (interactions between patient and treatment). Some features were based on expert knowledge or literature (e.g., classification of AED as first or second generation, AED activity in particular seizure types, mechanism of action [12–14]). To reduce the number of features and avoid over-fitting the training data, we used a standard feature selection process. First, we removed constant features (mode frequency > 0.99) and features with small standard deviations. Second, we removed features whose correlation to the outcome was not significant ($p > 0.05$). Finally, we removed features that were highly correlated to another feature ($R > 0.6$), where the feature removed was the one with lower correlation to the outcome. The selected features were based on the training set only, and when testing, only these selected features were calculated and used. We used the random forest algorithm [15], a state-of-the-art prediction model that outputs the majority vote of a multitude of decision trees because this algorithm outperformed other prediction algorithms (results not shown).

The performance of the model's predictions was estimated using receiver operator characteristic (ROC) analysis on the test set. This accuracy was compared with the accuracy of a baseline model, which was defined as a logistic regression on the number of treatment changes the patient had experienced in the 12 months prior to the index date.