



HHS Public Access

Author manuscript

Med Care. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:

Med Care. 2019 April ; 57(4): 318–324. doi:10.1097/MLR.0000000000001072.

Epilepsy Among Elderly Medicare Beneficiaries:

A validated approach to identify prevalent and incident epilepsy

Lidia M. V. R. Moura, MD, MPH^{*,†,‡}, Jason R. Smith, BA^{*}, Deborah Blacker, MD, ScD^{†,§}, Christine Vogeli, PhD^{||}, Lee H. Schwamm, MD^{*,‡}, Andrew J. Cole, MD^{*,‡}, Sonia Hernandez-Diaz, MD, DPH[†], and John Hsu, MD, MBA^{#,**}

* Department of Neurology, Massachusetts General Hospital, Boston, MA

† Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

‡ Department of Neurology, Harvard Medical School, Boston, MA

§ Department of Psychiatry, Massachusetts General Hospital, Boston, MA

|| Department of Medicine, Massachusetts General Hospital, Boston, MA

Mongan Institute, Department of Medicine, Massachusetts General Hospital, Boston, MA

** Department of Health Care Policy, Harvard Medical School, Boston, MA.

Abstract

Background: Uncertain validity of epilepsy diagnoses within health insurance claims and other large datasets have hindered efforts to study and monitor care at the population level.

Objectives: To develop and validate prediction models using longitudinal Medicare administrative data to identify patients with actual epilepsy amongst those with the diagnosis.

Research Design, Subjects, Measures: We used linked electronic health records (EHR) and Medicare administrative data including claims to predict epilepsy status. A neurologist reviewed EHR data to assess epilepsy status in a stratified random sample of Medicare beneficiaries aged 65+ years between 01/2012–12/2014. We then reconstructed the full sample using inverse

Corresponding Author: Lidia M. V. R. Moura, 55 Fruit St., Wang ACC 739D, Boston, MA 02114, Phone: 617-726-3311, Fax: 617-726-9250, lidia.moura@mgh.harvard.edu.

COMPLETE AUTHOR INFORMATION

Lidia M. V. R. Moura, MD, MPH; 55 Fruit St., Wang ACC 739D, Boston, MA 02114; Phone: 617-726-3311; Fax: 617-726-9250; lidia.moura@mgh.harvard.edu

Jason R. Smith, BA; 55 Fruit Street, Wang ACC 731A, Boston, MA 02114; Phone: 617-726-6896; Fax: 617-726-9250; jason.smith@mgh.harvard.edu

Deborah Blacker, MD, ScD; Neurology Associates, 15 Parkman Street, Boston, MA 02114; Phone: 617-726-1728; Fax: 617-726-4101; blacker@nmr.mgh.harvard.edu

Christine Vogeli, PhD; 55 Fruit Street, Boston, MA 02114; Phone: 617-724-0984; Fax: cvogeli@mgh.harvard.edu

Lee H. Schwamm, MD; Neurology & Stroke Services, 15 Parkman Street, Boston, MA 02114; Phone: 617-726-8459; Fax: 617-726-5043; lschwamm@mgh.harvard.edu

Andrew J. Cole, MD; Neurology Associates, 15 Parkman Street, Wang ACC 739L, Boston, MA 02114; Phone: 617-726-3311; Fax: 617-726-9250; cole.andrew@mgh.harvard.edu

Sonia Hernandez-Diaz, MD, DPH; Kresge Buidling, Room 816B, 677 Huntington Avenue, Boston, MA 02115; Phone: 617-432-3942; Fax: sherman@hsph.harvard.edu

John Hsu, MD, MBA; Institute for Health Policy, Massachusetts General Hospital, 50 Staniford Street, Boston, MA 02114; Phone: 617-643-7767; Fax: 617-724-4738; john.hsu@mgh.harvard.edu

Conflicts of Interest

None of the authors has any conflict of interest to declare in relation to this work.

probability sampling weights. We developed prediction models using longitudinal Medicare data, then in a separate sample evaluated the predictive performance of each model, e.g., area under the receiver operating characteristic curve (AUROC), sensitivity, and specificity.

Results: Of 20,945 patients in the reconstructed sample, 2.1% had confirmed epilepsy. The best performing prediction model to identify prevalent epilepsy required epilepsy diagnoses with multiple claims at least 60-days apart, and epilepsy-specific drug claims: AUROC = 0.93 (95% CI 0.90–0.96), and with an 80% diagnostic threshold, sensitivity = 87.8% (95% CI 80.4%–93.2%), specificity = 98.4% (95% CI 98.2%–98.5%). A similar model also performed well in predicting incident epilepsy ($k = 0.79$; 95% CI 0.66 – 0.92).

Conclusions: Prediction models using longitudinal Medicare data perform well in predicting incident and prevalent epilepsy status accurately.

Keywords

Epilepsy; epidemiology; elderly; claims data; algorithms

INTRODUCTION

Epilepsy is a life-threatening, often lifelong, disorder characterized by recurrent, spontaneous seizures.^{1, 2} Presentation of epilepsy can be varied, in part due to the range of etiologies. For example, epilepsy may arise from common structural pathologies including neurodegeneration, trauma, stroke, and tumors.^{3, 4} Epilepsy may present with subtle cognitive or behavioral symptoms, which could be mistaken for age-related cognitive impairment.^{5, 6} At times, both epilepsy and the seizures both go undetected.³ Early detection and treatment, however, is critical as the risk of seizure recurrence after a first seizure can be as high as 80%.^{7, 8} Moreover, morbidity can be exacerbated by other comorbidities associated with advanced age.⁹

Despite the significant health burden posed by epilepsy, there are few national estimates of disease incidence or prevalence, though anecdotal estimates suggest that both have increased. Generating such national estimates and tracking care for patients across the country would require a well validated approach to identify true epilepsy cases using data from large, national datasets, e.g., Medicare administrative datasets including insurance claims. Relying only on diagnosis codes, however, is a potentially fraught process.^{10–14} At the same time, many large datasets contain more information than only diagnosis codes, some of which could be useful for predicting disease status.

Furthermore, the diagnosis and treatment of epilepsy in the elderly requires nuances in clinical judgement and review of information that might not be captured in claims, e.g., synthesis of non-specific symptoms, detailed histories, and impressions from diagnostic tests. As a result, previous studies that ignore age and focus primarily on diagnostic and treatment criteria have failed to produce accurate and validated claims-based measures to detect epilepsy among individuals aged 65+ years.^{11, 12, 15–20} These methodological limitations contribute to the uncertainty about the true prevalence of epilepsy in the elderly population.

To address this gap in the literature, we used a population-based study design linking longitudinal Medicare administrative data (e.g., claims) to electronic health records (EHR) data, combined with blinded expert review of the clinical data, to determine whether claims-based models can accurately predict which patients in the elderly Medicare population have true epilepsy disease.

METHODS

Data Sources and Sampling Approach

We used three longitudinal datasets: 1) The Partners Healthcare System EHRs; 2) The Partners Accountable Care Organization (ACO) claims data. We used Medicare claims from Medicare Parts A/B (hospital and physician services) and D (prescription drugs). The Partners Healthcare Institutional Review Board approved the study protocol.

We used a three-year observation period (01/2012–12/2014), which is a reasonable length to capture an established diagnosis of epilepsy.^{2, 11, 14, 16} We then used the following eligibility criteria: 1) aged 65+ years by 01/2012; 2) enrollment in both Medicare Parts A and B; 3) continuous alignment to the ACO between 01/2012–12/2014 or until death; 4) Medicare original reason for entitlement code (OREC) of age or disability; and 5) community dwelling (not institutionalized) at the time of ACO entry.²¹ To improve the efficiency of the chart reviews, we then performed a structured sampling on the full sample (n = 20,945) based on claims coded under the International Classification of Diseases, ninth revision (ICD-9) (Figure 1). Table 1 provides descriptive characteristics of the three samples (total n=1,534): A) at least one claim potentially indicative of epilepsy (codes 345.xx; 780.39); B) no claim for epilepsy but seen by a neurologist; and C) no claims for epilepsy and not seen by a neurologist. Our rationale was three-fold: a) it increased the power for analysis involving between-group comparisons because analytical groups were of reasonably comparable sizes (i.e., true epilepsy cases versus not, as well as false positives and false negative cases), b) it increased the feasibility of the study given finite resources, and c) it enabled us to reconstruct the target population rates using the inverse of the probability of being sampled weights.

EHR-based Diagnosis (Reference Standard)

An experienced neurologist reviewed and abstracted clinical data from all 1,534 EHRs (including physician notes, emergency department visit records, and diagnostic tests such as EEG files or brain imaging), without information on which of the three samples the patient was a member or on any of the claims information. Using the current International League Against Epilepsy (ILAE) guidelines (Text, Supplemental Digital Content 1, and Table, Supplemental Digital Content 2), the neurologist determined prevalent epilepsy disease status and epilepsy classification.^{22, 23}

Within the sample of prevalent cases, the neurologist then determined which of these first occurred between 2012–14, i.e., potential incident cases. The neurologist obtained the year of first epilepsy diagnosis (i.e., the index event) and the year of the first documented seizure. The neurologist exclusively designated potential incident epilepsy cases after careful review

of physician notes and records, conservatively excluding patients with documentation of a definitively acknowledged epilepsy diagnosis or a potential seizure before 2012.

Claims Diagnostic Codes

The goal of model development was to predict the probability of having epilepsy using the longitudinal administrative Medicare data, including the Medicare claims. To this end, we first used expert clinical knowledge to create lists of codes suggestive of epilepsy care (provided in the Table, Supplemental Digital Content 3). Next, we combined varying criteria (such as occurrence, position, or counts) to create several ordinal candidate variables (Table 2). Some ordinal variables could increase the probability of epilepsy, e.g., two codes for epilepsy as a primary, secondary, or tertiary diagnosis 60 days apart. Conversely, others might decrease the probability, e.g., diagnosis for recurrent syncope, as they may trigger work-up for loss of awareness that may lead to miscoding for epilepsy. We constructed these variables to reflect the potential diversity in epilepsy care, as patients can receive care and treatment at varying levels (such as primary care physician offices, general neurology clinics, or epilepsy sub-specialty clinics). We optimized variable performance by accounting for varying levels of granularity in these discrete encounters.

Finally, we used inverse probability sampling weights to recreate the population (i.e., we used the EHR review sample to efficiently reconstruct the full sample but with EHR-based epilepsy diagnoses). We accounted for age and gender and analyzed the distribution of clinical characteristics with confirmed epilepsy and non-confirmed epilepsy.

Model Development

To develop and validate models in a split-sample approach we randomly divided the reconstructed sample into two subsets: 1) model development dataset (n = 10,518; 50.9%), which included the indicator variable for EHR-based diagnosis for each case; and 2) validation dataset (n = 10,139; 49.1%), in which the indicator variable for EHR-based diagnosis was removed. We checked covariate balance to assess the adequacy of the randomization.

We first developed several multivariable logistic models to predict the probability of reference-standard prevalent epilepsy. We used clinical judgement to choose one variable indicative of an epilepsy diagnosis (“epilepsy diagnosis”, Table 2). “Epilepsy diagnosis” is an ordinal variable with “weights” ranging from 0 to 4, that adds one weight for the presence of ICD-9 claims coded as primary, secondary or tertiary diagnosis of epilepsy, convulsions, syncope, or collapse (i.e., three weights total if at least one claim in the primary, secondary, and tertiary diagnosis) and gives an extra weight for the presence of more than one primary epilepsy claim at least 60 days apart (i.e., maximum of four “weights”).

We also included variables that could represent potential false-positives (i.e., alternative conditions on the list of differential diagnoses). Several of these variables were tested in varying combinations to evaluate initial model performance (measured by AUROC). We also included variables that reflected procedures performed, specialty of the physician making the diagnosis, and/or site of care variables. To conclude, we added treatment variables: 1) “epilepsy-specific AED” (Table 2), which functioned by assigning a weight of 1 for the

presence of at least one epilepsy-specific drug (i.e., lamotrigine, levetiracetam, valproic acid, phenytoin, and carbamazepine), and a weight of 0 if otherwise; or 2) “AED” (Table 2), which functioned exactly as epilepsy-specific AED, but included all AEDs (i.e., AEDs that, after expert clinical review, are more frequently prescribed for indications aside from epilepsy, including gabapentin, topiramate, benzodiazepines). Different models used different epilepsy-related covariates to reflect varying potential uses (e.g., some researchers might have access to diagnosis and drug data, others might only have access to diagnosis data).

Statistical Analysis and Model Validation

Using the validation dataset, we assessed the performance of the different prediction models. Our performance measures included sensitivity (i.e., the percentage of total epilepsy cases correctly identified as epilepsy cases by the model), specificity (i.e., the percentage of total non-epilepsy cases correctly identified as non-epilepsy by the model), and the area under the receiver operating curve or AUROC (i.e., a measure of classifying a true epilepsy case higher than a false epilepsy case). We also calculated the positive predictive values (PPV; i.e., model-identified “true” epilepsy cases) and negative predictive values (NPV; i.e., model-identified “true” non-epilepsy cases), which by definition are sample-specific measures that vary with the underlying prevalence. While we report all of the performance measures and they tended to track together, we primarily use the AUROC as the most relevant measure.^{24–26} We include additional details on statistical code in Supplemental Digital Content 1.

We conducted several further analyses on the best performing prediction model. We first altered the diagnosis threshold (i.e., we increased, then decreased, the threshold for an individual to be classified as having epilepsy by the model) to assess the impact on sensitivity and specificity. Next, to imitate conditions in which less data might be available, we examined predictive power after restricting claims-based variables to a one-year window (01/2012–12/2012). And third, we examined performance after conservatively excluding reference-standard non-epilepsy cases due to “incomplete chart but no evidence to indicate epilepsy” (n = 181).

Finally, among those patients with epilepsy identified by the best performing model (i.e., prevalent epilepsy between 01/2012–12/2014), we analyzed the 3-year incidence rate using two methods. In the first method, we applied washout criteria (no claim indicative of epilepsy and no epilepsy-specific AED prescription) for periods of 6, 12, 18, 24, 36, and 48 months prior to the index event.²⁷ In the second method, we applied washout criteria exclusively to 01/2009–12/2011. The best-performing method was chosen by its level of agreement with diagnosis dates extracted from review of EHRs, measured by Cohens kappa coefficient.

Sensitivity Analysis

We conducted three types of sensitivity analyses. First, because more severely ill patients, or patients seen by a neurologist, may yield richer (or more accurate) epilepsy care documentation, we performed stratified analyses on the best performing model. We separately analyzed the patients seen by a neurologist (n = 6,785) and those without

neurologist involvement (n = 16,160). Second, we examined the accuracy of two previously published algorithms^{2, 18} when applied to our reconstructed sample of Medicare beneficiaries to evaluate accuracy and generalizability: 1) one claim for epilepsy as primary diagnosis (345.xx) or at least two claims for convulsions (789.3x) at least 30 days apart²; 2) diagnosis codes potentially indicative of epilepsy and AED prescriptions¹⁸. The variables we constructed to recreate these algorithms are detailed in Table 2. And third, because epilepsy is not always diagnosed concurrently with a first seizure diagnosis, we analyzed the performance of our methods to measure 3-year incidence rates by altering the index event of the date of first diagnosis to “date of first seizure” and reporting agreement with EHR-review (measured by kappa coefficient).

RESULTS

Patient Characteristics

In the reconstructed sample, 2.1% or 21/1,000 patients had a reference standard epilepsy diagnosis over the 3-year period (Figure 1, Table 1). Of those patients with prevalent epilepsy (mean age = 76.4 years [SD = 8.5 years]), 230 (52.9%) were female and 205 (47.1%) were male. We provide descriptive information about each EHR review stratum in the Supplemental Digital Content 4 (Text) and Supplemental Digital Content 5–7 (Tables).

Claims Diagnostic Codes

We report the distribution of all claims-based variables stratified by epilepsy diagnosis (reference standard) in the table of Supplemental Digital Content 8. For instance, all suspected claims-based epilepsy variables (e.g. epilepsy diagnosis, primary epilepsy diagnosis) occurred more often among patients with epilepsy compared to no epilepsy (e.g. primary epilepsy diagnosis: 59.8% and 13.5%, respectively, $p < 0.0001$).

Models

Epilepsy diagnosis was evenly and randomly distributed across the model development and model validation subset: 10,336 patients with no epilepsy (51.1%) were distributed in the development subset and 9,886 (48.9%) were distributed into the validation subset; 182 patients with epilepsy (41.8%) were randomly distributed into the development subset and 253 (58.2%) into the validation subset.

Three models were chosen based on their ability to predict the neurologist’s EHR-based diagnosis of epilepsy (Table 3, Table 4). The best performing model (model 1), defined by the highest AUROC, included: an epilepsy diagnosis code variable and epilepsy-specific AED variable (sensitivity = 87.8%, 95% CI 80.4%–93.2%; specificity = 98.4%, 95% CI 98.2%–98.5%; AUROC = 0.93, 95% CI 0.90–0.96; PPV = 23.1%, 95% CI 19.2%–27.4%). Increasing the diagnosis threshold of model 1 from 0.80 to 0.90 maximized sensitivity and AUROC over specificity (sensitivity = 100.0%, 95% CI 93.5%–100.0%; specificity = 98.2%, 95% CI 98.0%–98.3%; AUROC = 0.99, 95% CI 0.99–0.99; PPV = 12.6%, 95% CI 9.6%–16.1%; NPV = 100.0, 95% CI 100.0%–100.0%), meanwhile decreasing the diagnosis threshold to 0.05 maximized specificity and PPV over sensitivity (sensitivity = 18.3%, 95% CI 16.5%–20.3%; specificity = 99.3%, 95% CI 99.2%–99.4%; AUROC = 0.59, 95% CI

0.58–0.60; PPV = 69.6%, 95% CI 65.0%–73.8%; NPV = 93.3%, 95% CI 93.0%–93.7%). This indicates that when constructing prediction models, different diagnosis thresholds should be examined to achieve balanced sensitivity and specificity.

Of those patients with epilepsy identified by model 1 ($n = 114$), the most accurate method to identify incident epilepsy was to apply washout criteria exclusively to 01/2009–12/2011 ($k = 0.79$, 95% CI 0.66–0.92). The other tested approach (i.e., washout periods of no claims indicative of epilepsy and no epilepsy-specific AED claims directly before the identified index event) for substantiating an incident epilepsy case resulted in a severe drop to the measure of agreement. A 48-month washout period resulted in a k coefficient of 0.02 (95% CI –0.00 – 0.05); 6 months: $k = -0.30$ (95% CI –0.38 - –0.21); 12 months: $k = -0.38$ (95% CI –0.52 - –0.24); 18 months: $k = -0.47$ (–0.63 - –0.31); 24 months: $k = -0.49$ (95% CI –0.66 - –0.32); and 36 months: $k = -0.52$ (95% CI –0.71 - –0.34).

Supplemental Digital Content 4 (Text) and Supplemental Digital Content 9 (Table) provide further information on two additional high-performing models, and further analyses on model 1.

Sensitivity Analysis

Models derived from previously published algorithms^{2, 18} performed poorly in our reconstructed sample. One approach² maximized specificity at the cost of diminished sensitivity (sensitivity = 2.7%, 95% CI 2.5%–3.0%; specificity = 100.0%, 95% CI 99.9%–100.0%; PPV = 100.0%, 95% CI 99.2%–100.0%; NPV = 22.6%, 95% CI 22.1%–23.2%; AUROC = 0.51, 95% CI 0.51–0.51). Another approach¹⁸, with which we set a conservative diagnosis threshold of $p > 0.10$, likewise maximized specificity and PPV over sensitivity and AUROC (sensitivity = 17.3%, 95% CI 15.1%–19.8%; specificity = 98.7%, 95% CI 98.5%–98.8%; PPV = 40.2%, 95% CI 35.6%–44.9%; NPV = 95.9%, 95% CI 95.6%–96.1%; AUROC = 0.58, 95% CI 0.57–0.59).

Analyzing incident epilepsy cases using the date of first seizure diagnosis yielded results consistent with the date of first epilepsy diagnosis. Using the method of no claims between 01/2009–12/2011, $k = 0.76$ (95% CI 0.62 – 0.90); 48-month washout period, $k = 0.02$ (95% CI –0.00 – 0.04); 6 months, $k = -0.29$ (95% CI –0.37 - –0.21); 12 months, $k = -0.36$ (95% CI –0.50 - –0.22); 18 months, $k = -0.44$ (95% CI –0.60 - –0.29); 24 months, $k = 0.47$ (95% CI –0.63 - –0.30); and 36 months, $k = -0.50$ (95% CI –0.68 - –0.32). Supplemental Digital Content 4 (Text) supplies final sensitivity analysis on model 1, where we stratify our sample by neurology specialty involvement.

DISCUSSION

In the first rigorous validation of models for epilepsy identification in the elderly using large datasets, we found that longitudinal Medicare administrative data can accurately detect prevalent and incident epilepsy using a multivariate prediction modeling approach. The most accurate model required a claims-based diagnosis code for epilepsy and evidence of epilepsy-specific drug use. Incorporating other relevant clinical data such as differential diagnosis and procedure data yielded similarly high levels of accuracy but were arguably

more complex. In short, the general prediction approach and range of potential methods illustrate multiple viable options that others could match with their needs and data contexts (e.g., surveillance studies with limited access to certain data).

The most recent estimates of epilepsy incidence and prevalence in the elderly population using Medicare administrative data in the US revealed an annual incidence of 6.1/1,000 individuals and prevalence of 15.2/1,000 individuals.¹⁴ This study utilized an un-validated algorithm that, when applied to a nationwide sample of Medicare beneficiaries, yielded an annual prevalence of 10.8/1,000.² Our study demonstrated this approach² strongly favored specificity over sensitivity (specificity = 100%, 95% CI 99.9%–100%; sensitivity = 2.7%, 95% CI 2.5%–3.0%) when applied to our dataset, which suggests this algorithmic approach is not versatile to identify individuals with epilepsy in outside datasets and might not be suited for tracking patient care nationwide.

Prior algorithms have reported moderate performance albeit often requiring difficult tradeoffs. One recent study used data from the health information services department at a hospital in Melbourne, Australia to validate an algorithm utilizing ICD-10 codes for epilepsy and 1 AED, favoring high specificity at the cost of sensitivity (sensitivity = 60.8%, specificity = 99.9%, PPV = 81.4%).¹⁷ Conversely, others reported high sensitivity (85%–95%) and specificity (87%–99%), but their models performed poorly when applied to a second dataset, suggesting overfitting.^{11, 12, 15, 16} Our predictive models were developed and validated in a split-sample approach, and attained high accuracy for the function of identifying epilepsy.

Yet, there are a host of additional difficulties that belie simple epilepsy and non-epileptic event identification in the elderly. Of note, pre-existing dementia, for example, can obscure epilepsy symptoms (e.g. pre-existing confusional states, delirium). At the same time, events common in the elderly population (e.g. syncope, transient ischemic events) often mimic symptoms of epilepsy and widen the differential diagnoses. Compared to prior validation studies that reviewed and assessed epilepsy diagnosis via non-specialist physicians or researchers,^{12, 28} our reference standard defined by a neurologist decreased the chance for erroneous assessment.

Further compounding the problem of diagnosis and treatment is difficulty in differentiating between non-epileptic and epileptic seizures based on a single claim. Our review of EHRs, which yielded an accurate portrayal of the elderly population with epilepsy, elucidated that non-epileptic events (i.e., events such as syncope or altered mental status) are common among the elderly population (e.g., 987 patients (4.9%) were diagnosed with a non-epileptic event), and these, in turn, represented most of the claims-based false positives. The data is consistent with clinical practice and prior studies demonstrating that the differentiation between epileptic and non-epileptic events remains a challenge, particularly in the elderly.²⁹

Limitations

We designed our methods of sampling and data collection with prevalence in mind. Prediction of disease onset and development of future disease merit attention in future work. Similarly, our strategy of using claims data over a three-year period for disease status

prediction also relies on the diagnostic and treatment patterns within the US Medicare population aged 65 years or older. Thus, our findings have unclear generalizability to populations in other countries or in younger populations, which might have different care patterns.

Additionally, sampling patients enrolled in the ACO may limit generalizability. These patients could have a higher (or lower) level of education and/or more severe disease (i.e. a continuous need for Medicare enrollment); thus, patients with specific levels of education and/or less severe disease might be underrepresented. Although we did not have access to data on patient race or socioeconomic status through the ACO, we did document a broad range of both epilepsy syndromes/etiologies and differential diagnoses through EHR review (Tables, Supplemental Digital Content 5 and 6, which provide a breakdown of epilepsy syndrome/etiologies and differential diagnoses of patients in the reconstructed sample) (but we acknowledge that we are unable to assess the distribution of these factors).

Moreover, our study had a finite number of additional discriminatory analyses that we could feasibly perform. Although we had access to a large sample of patients spanning several hospitals and health centers, we did not interrogate the consistency of our model's performance across each discrete center. Further, we did not assess how accurately ICD-9 claims data captured the specific epilepsy syndrome, etiology, seizure type, or differential diagnosis (i.e., sub-decimal coding). The very limited studies that do classify epilepsy patients by etiologies based on claims still require validation.^{27, 30} Future validation is critical in order to provide accurate reports of epilepsy prevalence by etiology and syndromes,³¹ and there is the potential to increase the accuracy of identifying non-epileptic events in the elderly to promote more nuanced treatment.

Finally, we used only one neurologist reviewer from a leading neurology department in this study. Due to potential inter-reviewer variation in determining epilepsy status, reproducibility of our findings might be limited.

CONCLUSIONS

In this validation study, prediction models using longitudinal data from large administrative datasets were able to identify epilepsy status accurately. These types of large dataset electronic disease signatures represent the first step in monitoring epilepsy nationally and gathering real-world evidence on care for patients with epilepsy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank biostatistician Hang Lee, PhD of Massachusetts General Hospital, who revised statistical analysis through the Harvard Catalyst statistical consulting program.

Disclosure of Funding

This study was supported by the NIH (1K08AG053380-01A1) and the Harvard University Diversity Inclusion and Community Partnership program for protected research time to design and conduct all aspects of the study. Lidia M.V.R. Moura receives funding from the NIH (1K08AG053380-01A1), the Harvard University Diversity Inclusion and Community Partnership Faculty Fellowship Award, and the Schwamm-Marriott Clinical Care Research Fellowship Award. Deborah Blacker receives funding from the NIH (P50 AG005134, 1P01AG036694, U01 AG032984, U01AG045390). Sonia Hernandez-Diaz receives funding from the NIH (R21HD092879). John Hsu receives funding from the NIH (1R01 CA164023-04, 2P01AG032952-06A1, R01 HD075121-04, R01 MH104560-02).

References

1. Stovner LJ, Hoff JM, Svalheim S, Gilhus NE. Neurological disorders in the Global Burden of Disease 2010 study. *Acta Neurol Scand Suppl* 2014;1–6. [PubMed: 24588499]
2. Faught E, Richman J, Martin R, et al. Incidence and prevalence of epilepsy among older U.S. Medicare beneficiaries. *Neurology* 2012;78:448–453. [PubMed: 22262750]
3. Rowan AJ, Ramsay RE, Collins JF, et al. New onset geriatric epilepsy: a randomized study of gabapentin, lamotrigine, and carbamazepine. *Neurology* 2005;64:1868–1873. [PubMed: 15955935]
4. Leppik IE. *Epilepsy in the Elderly: Scope of the Problem International Review of Neurobiology*; Academic Press, 2007: 1–14.
5. Silveira DC, Jehi L, Chapin J, et al. Seizure semiology and aging. *Epilepsy Behav* 2011;20:375–377. [PubMed: 21273137]
6. Morillo LE. Temporal lobe epilepsy in the elderly. *Epilepsy Res Treat* 2012;2012:641323. [PubMed: 22957236]
7. Ramsay RE, Rowan AJ, Pryor FM. Special considerations in treating the elderly patient with epilepsy. *Neurology* 2004;62:S24–S29. [PubMed: 15007161]
8. Berg AT. Risk of recurrence after a first unprovoked seizure. *Epilepsia* 2008;49 Suppl 1:13–18.
9. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *The New England Journal of Medicine* 1988;1701. [PubMed: 3205267]
10. Tsiropoulos I, Gichangi A, Andersen M, Bjerrum L, Gaist D, Hallas J. Trends in utilization of antiepileptic drugs in Denmark. *Acta Neurol Scand* 2006;113:405–411. [PubMed: 16674607]
11. Reid AY, St Germaine-Smith C, Liu M, et al. Development and validation of a case definition for epilepsy for use with administrative health data. *Epilepsy Res* 2012;102:173–179. [PubMed: 22727659]
12. Franchi C, Giussani G, Messina P, et al. Validation of healthcare administrative data for the diagnosis of epilepsy. *J Epidemiol Community Health* 2013;67:1019–1024. [PubMed: 24022813]
13. Betjemann JP, Josephson SA, Lowenstein DH, Burke JF. Trends in Status Epilepticus-Related Hospitalizations and Mortality: Redefined in US Practice Over Time. *JAMA Neurol* 2015;72:650–655. [PubMed: 25915004]
14. Ip Q, Malone DC, Chong J, Harris RB, Labiner DM. An update on the prevalence and incidence of epilepsy among older adults. *Epilepsy Research* 2018;139:107–112. [PubMed: 29220741]
15. Moura LMVR, Price M, Cole AJ, Hoch DB, Hsu J. Accuracy of claims-based algorithms for epilepsy research: Revealing the unseen performance of claims-based studies. *Epilepsia* 2017;58:683–691. [PubMed: 28199007]
16. Jette N, Reid AY, Quan H, Hill MD, Wiebe S. How accurate is ICD coding for epilepsy? *Epilepsia* 2010;51:62–69. [PubMed: 19682027]
17. Tan M, Wilson I, Braganza V, et al. Development and validation of an epidemiologic case definition of epilepsy for use with routinely collected Australian health data. *Epilepsy Behav* 2015;51:65–72. [PubMed: 26262935]
18. Holden EW, Thanh Nguyen H, Grossman E, et al. Estimating prevalence, incidence, and disease-related mortality for patients with epilepsy in managed care organizations. *Epilepsia* 2005;46:311–319. [PubMed: 15679513]
19. Wassenaar M, Carpay JA, Sander JW, Thijs RD. Validity of health insurance data to identify people with epilepsy. *Epilepsy Research* 2018;139:102–106. [PubMed: 29220740]
20. Choi H, Pack A, Elkind MSV, et al. Predictors of incident epilepsy in older adults: The Cardiovascular Health Study. *Neurology* 2017;88:870–877. [PubMed: 28130470]

21. Yun H, Kilgore ML, Curtis JR, et al. Identifying types of nursing facility stays using medicare claims data: an algorithm and validation. *Health Services and Outcomes Research Methodology* 2010;10:100–110.
22. 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;58:512–521. [PubMed: 28276062]
23. 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;58:512–521. [PubMed: 28276062]
24. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. *Journal of clinical epidemiology* 2008;61:1085–1094. [PubMed: 19208371]
25. Riley RD, Ensor J, Snell KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ (Clinical research ed)* 2016;353:i3140.
26. Debray TP, Damen JA, Snell KI, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ (Clinical research ed)* 2017;356:i6460.
27. Thurman DJ, Beghi E, Begley CE, et al. Standards for epidemiologic studies and surveillance of epilepsy. *Epilepsia* 2011;52 Suppl 7:2–26. [PubMed: 21899536]
28. Fonferko-Shadrach B, Lacey AS, White CP, et al. Validating epilepsy diagnoses in routinely collected data. *Seizure* 2017;52:195–198. [PubMed: 29059611]
29. Leppik IE. Epilepsy in the Elderly: Scope of the Problem. *The Neurobiology of Epilepsy and Aging* 2007: 1–14.
30. Kurth T, Lewis BE, Walker AM. Health care resource utilization in patients with active epilepsy. *Epilepsia* 2010;51:874–882. [PubMed: 20002149]
31. Fiest KM, Sauro KM, Wiebe S, et al. Prevalence and incidence of epilepsy. *Neurology* 2017;88:296–303. [PubMed: 27986877]

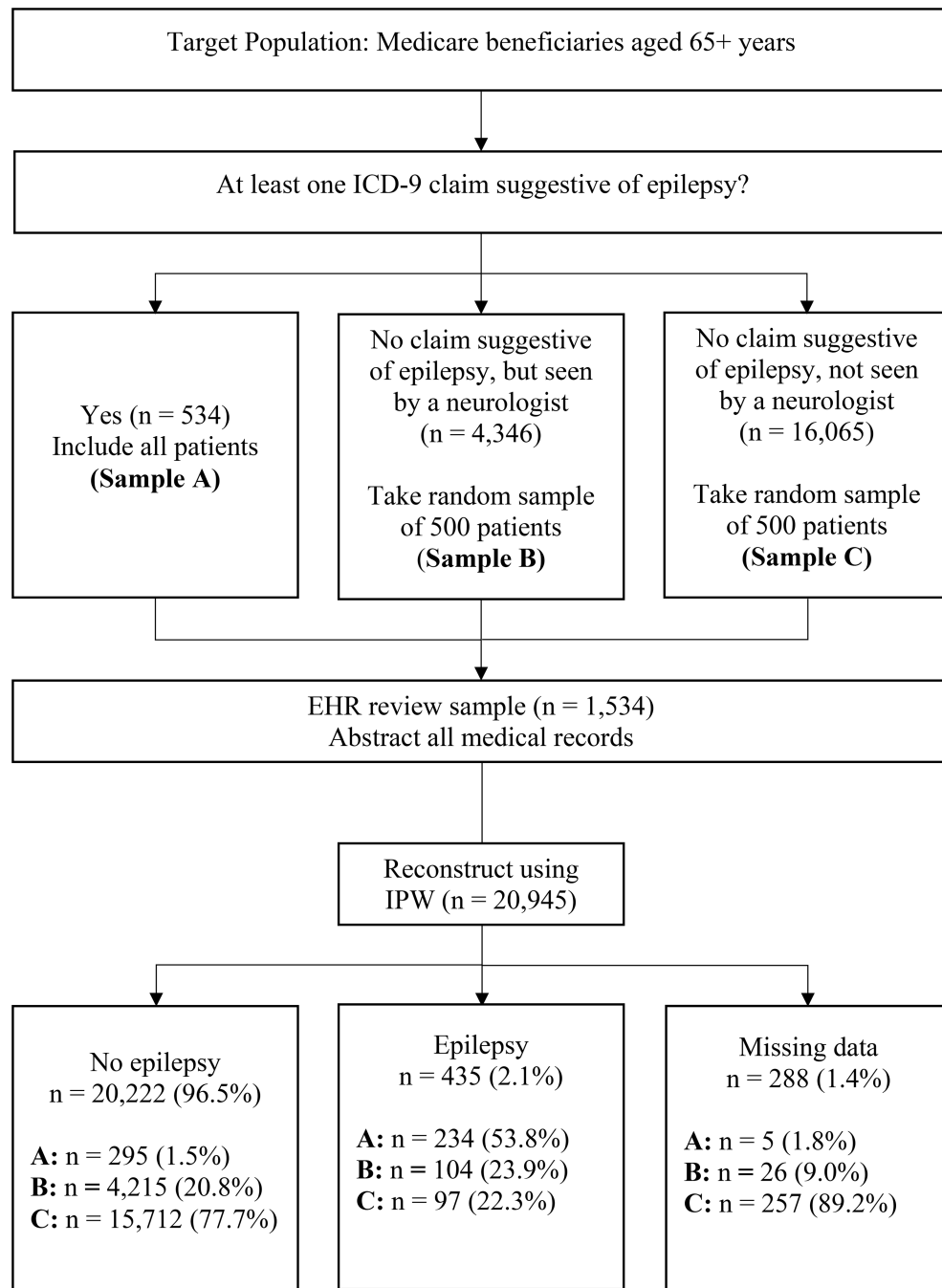


Figure 1. Sampling strategy and patients with adjudicated epilepsy diagnosis

Abbreviations: ICD-9 = International Classification of Diseases, Ninth Revision; IPW = inverse probability weighting.

Sampling strategy of eligible patients that resulted in a study sample of 1,534 patients, broken up into three sub-samples. ICD-9 claims related to epilepsy, the basis of stratification, were restricted to the time frame of 01/2012-12/2014.

Table 1.

Characteristics of reconstructed sample

Characteristic	Sample A At least one epilepsy claim (n = 534)	Sample B No epilepsy claims and not seen by neurologist (n = 16,065)	Sample C No epilepsy claims but seen by neurologist (n = 4,346)
Demographic ^a			
Age, mean (SD)	77.3 (7.2)	75.6 (39.4)	73.7 (19.1)
Female, n (%)	322 (60.3)	9,446 (58.8)	2,582 (59.4)
Treatment ^b			
Epilepsy-specific AED, n (%)	183 (34.3)	161 (1.0)	235 (5.4)
AED, n (%)	271 (50.7)	4,016 (25.0)	1,660 (38.2)
Suspected diagnosis ^a			
No epilepsy, n (%)	295 (55.2)	15,712 (97.8)	4,215 (97.0)
Epilepsy, n (%)	234 (43.8)	97 (0.6)	104 (2.0)
Missing data, n (%)	5 (1.0)	257 (1.6)	26 (1.0)

Abbreviations: AED = antiepileptic drug; ICD-9 = International Classification of Diseases, Ninth Revision.

^a information abstracted from electronic health records

^b information derived from Medicare claims.

Characteristics of reconstructed sample (n = 20,945) following electronic health record review. Missing data (n = 288) were reported when there was incomplete patient EHR information to ascertain an epilepsy diagnosis.

Table 2.

Candidate variables constructed from claims

Variable	Description
Suspected epilepsy ^a	
Primary epilepsy diagnosis	Epilepsy as a primary diagnosis
Epilepsy diagnosis ^b	Weighted; Epilepsy as primary, secondary, or tertiary diagnosis, and gives an additional weight to the occurrence of two or more primary codes 60 days apart
Algorithmic epilepsy diagnosis 1 ^c	Any epilepsy diagnosis
Algorithmic epilepsy diagnosis 2 ^c	Convulsions at least 30 days apart coinciding with inpatient stays, outpatient visits, or physician visits within one year
Potential false positive ^a	
Withdrawal seizure	Diagnosis of seizure due to alcohol or drug withdrawal
Drug reaction seizure	Diagnosis of seizure due to a drug reaction
Syncope and collapse event	Diagnosis of "syncope or collapse"
Alteration of consciousness event	Diagnosis of alteration of consciousness event
Sleep disturbance event	Diagnosis of sleep disturbance event
Dementia	Diagnosis of dementia
Migraine event	Diagnosis of migraine event
Psychogenic event	Diagnosis of psychogenic event
Cerebrovascular event	Diagnosis of cerebrovascular event or cognitive deficit
Non-epileptic event diagnosis	Potential false positive diagnosis, excluding withdrawal and drug reaction seizures
Non-epileptic event or seizure diagnosis	Any potential false positive diagnosis
Specialty involvement ^d	
Neurology	Epilepsy-related claim indicative of neurology specialty involvement
Site of care ^d	
Urgent care and ER	Epilepsy-related claim indicative of service provided at an urgent care facility or hospital ER
Procedure ^e	
Brain imaging	Epilepsy-related claim indicative of a brain imaging procedure
EEG	Epilepsy-related claim indicative of an EEG procedure
Treatment	
Epilepsy-specific AED	Specific AED prescription (levetiracetam, lamotrigine, valproic acid, phenytoin, or carbamazepine) with a concurrent epilepsy diagnosis code
AED ^b	All AED prescriptions

Abbreviations: ER = emergency room; EEG = electroencephalography; AED = antiepileptic drug; ICD-9 = International Classification of Diseases, Ninth Revision; CPT = current procedural terminology.

^aICD-9 codes

^bvariables used to recreate Holden et al algorithm¹

^cvariables created based on Faught et al. algorithm²

d
service codes

e
CPT codes

List of variables constructed from ICD-9, CPT, and service codes for creation of models. Codes used to create each variable are listed in Supplemental Digital Content 2 (Table). Medicare claims were pulled between the dates 01/2012 and 12/2014.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Characteristics of highest performing claims-based models

Characteristic	Model 1 ^a	Model 2 ^b	Model 3 ^c
Suspected epilepsy			
Epilepsy diagnosis	Yes	Yes	Yes
Potential false positive			
Non-epileptic event or seizure diagnosis	No	Yes	Yes
Site of care			
Urgent care and ER	No	Yes	Yes
Procedure			
EEG	No	Yes	Yes
Brain imaging	No	Yes	Yes
Specialty involvement			
Neurology	No	Yes	Yes
Treatment			
Epilepsy-specific AED	Yes	Yes	No

Abbreviations: ER = emergency room = EEG, electroencephalography; AED = antiepileptic drug.

^aModel 1: $\text{Logit}(\text{epilepsy}) = \beta_0 + \beta_1 * (\text{epilepsy diagnosis}) + \beta_2 * (\text{epilepsy-specific AED})$; where $\beta_0 = -3.1$, $\beta_1 = 0.9$, and $\beta_2 = 1.9$; number of observations read = 1,518; number of observations used = 1,516; c-statistic: 0.927

^bModel 2: $\text{Logit}(\text{epilepsy}) = \beta_0 + \beta_1 * (\text{epilepsy diagnosis}) + \beta_2 * (\text{epilepsy-specific AED}) + \beta_3 * (\text{alteration of consciousness}) + \beta_4 * (\text{sleep disturbance}) + \beta_5 * (\text{dementia}) + \beta_6 * (\text{non-epileptic event or seizure diagnosis}) + \beta_7 * (\text{urgent care and emergency room}) + \beta_8 * (\text{EEG}) + \beta_9 * (\text{brain imaging}) + \beta_{10} * (\text{neurology})$; detailed statistical output values omitted from this footnote.

^cModel 3: $\text{Logit}(\text{epilepsy}) = \beta_0 + \beta_1 * (\text{epilepsy diagnosis}) + \beta_2 * (\text{alteration of consciousness}) + \beta_3 * (\text{sleep disturbance}) + \beta_4 * (\text{dementia}) + \beta_5 * (\text{non-epileptic event or seizure diagnosis}) + \beta_6 * (\text{urgent care and emergency room}) + \beta_7 * (\text{EEG}) + \beta_8 * (\text{brain imaging}) + \beta_9 * (\text{neurology})$.

Claims-based variables used to construct best performing models.

Table 4.

Performance metrics of highest performing claims-based models

Model	AUROC (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Diagnosis threshold
Model 1	0.93 (0.90 – 0.96)	87.8 (80.4 – 93.2)	98.4 (98.2 – 98.5)	23.1 (19.2 – 27.4)	99.9 (99.9 – 100.0)	> 0.80
Model 2	0.91 (0.88 – 0.94)	83.9 (76.0 – 90.0)	98.4 (98.2 – 98.5)	22.6 (18.8 – 26.8)	99.9 (99.9 – 99.9)	> 0.80
Model 3	0.80 (0.72 – 0.88)	80.0 (59.3 – 93.2)	79.5 (77.1 – 81.7)	7.4 (4.6 – 11.2)	99.5 (98.8 – 99.8)	> 0.80

Abbreviations: AUROC = area under the receiver operating characteristic curve; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value.

Performance metrics of best performing models, determined by the AUROC.