Validation of a suspicion index to identify patients at risk for hereditary angioedema

Marissa Shams, MD,[a] Dawn A. Laney, MS,[a] Dave A. Jacob, BS,[b] Jingjing Yang, PhD,[a] Jessica Dronen, MS,[c] Amanda Logue, MD,[d] Ami Rosen, MS,[a] and Marc Riedl, MD[e] *Atlanta, Ga; Sudbury, Mass; Lafayette, La; and San Diego, Calif*

Background: Hereditary angioedema (HAE) is a genetic condition characterized by dysregulation of the contact (kallikrein-bradykinin) pathway, leading to recurrent episodes of angioedema.

Objective: This project sought to determine whether a suspicion index screening tool using electronic health record (EHR) data can identify patients with an increased likelihood of a diagnosis of HAE.

Methods: A suspicion index screening tool for HAE was created and validated by using known patients with HAE from the medical literature as well as positive and negative controls from HAE-focused centers. Through the use of key features of medical and family history, a series of logistic regression models for 5 known genetic causes of HAE were created. Top variables populated the digital suspicion scoring system and were run against deidentified EHR data. Patients at 2 diverse sites were categorized as being at increased, possible, or no increased risk of HAE.

Results: Prediction scoring using the strongest 13 variables on the "real-world" EHR-positive control data identified all but 1 patient with C1 inhibitor deficiency and patient with non–C1 inhibitor deficiency without false-positive results. The 2 missed patients had no documented family history of HAE in their EHR. When the prediction scoring variables were expanded to 25, the screening algorithm approached 100% sensitivity and specificity. The 25-variable algorithm run on general population EHR data identified 26 patients at the medical centers as being at increased risk for HAE.

Conclusions: These results suggest that development, validation, and implementation of suspicion index screening tools can be useful to aid providers in identifying patients with rare genetic conditions. (J Allergy Clin Immunol Global 2023;2:76-8.)

Key words: *Hereditary angioedema, suspicion index screening tool, electronic health record, genetics, angioedema*

INTRODUCTION

Hereditary angioedema (HAE) is a genetic condition characterized by dysregulation of the contact (kallikrein-bradykinin) pathway as well as the complement and coagulation systems, leading to recurrent swelling episodes (angioedema).[1] Angioedema commonly affects the digestive tract and extremities, but it can also affect the lips, tongue, face, genitals, and airway (larynx). Edema can be life-threatening, as it can occur unexpectedly and does not respond effectively to medications for anaphylaxis or mast cell–mediated angioedema.[1]

Many patients with HAE globally still experience long diagnostic delays, with a median of 6.5 years from symptom onset to diagnosis.[1-4] To reduce diagnostic delay, we developed, validated, and completed a pilot run of an automated suspicion index screening (SIS) tool using electronic health record (EHR) data. The aim of the tool is to provide an "early warning system" that identifies patients who are at risk for HAE and require further evaluation to maximize effective diagnosis and treatment. This clinical communication describes the use of real-world EHR data analysis to determine the discriminatory power of the SIS tool.

METHODS

The project began with HAE experts determining which data points they predicted to be most effective in identifying patients. Clinical variables considered as covariates were selected according to the following rules: (1) having nonzero variation among the patient cases and control groups, (2) showing no genetic information, and (3) having missing values treated as "absence" of the corresponding variable. All variables were binary, with 0 denoting absence and 1 denoting presence. A continuous age variable was also included in the logistic regression model as a covariate.

The positive control included patients with HAE caused by pathogenic variants in the *SERPING1* (n = 33), *PLG* (n = 20), *KNG1* (n = 8), *ANGPT1* (n = 1), and *F12* (n = 6) genes. The negative control included patients with lupus, medication-associated edema (excluding angiotensin-

From [a]the Emory School of Medicine, Atlanta; [b]the ThinkGenetic Foundation, Sudbury; [c]ThinkGenetic, Inc, Lafayette; [d]the Ochsner Lafayette General, Lafayette; and [e]the University of California, San Diego.

Supported in part by an investigator-initiated research grant from Takeda to the ThinkGenetic Foundation.

Western Institutional Review Board ethics committee approval no: 20200556.

Disclosure of potential conflict of interest: D. A. Laney is a consultant for and has received grants from Amicus Therapeutics, Chiesi, Protalix BioTherapeutics, Sanofi-Genzyme, Spark Therapeutics, and Takeda and a cofounder of ThinkGenetic, Inc; her activities have been monitored and found to be in compliance with the conflict of interest policies at the Emory University School of Medicine. D.A. Jacob is president of ThinkGenetic, Inc. J. Dronen is an employee of ThinkGenetic, Inc. M. A. Riedl is a consultant for Adverum, Attune, BioCryst, CSL Behring, KalVista, Pharming, Pharvaris, and Takeda and has received grants from BioCryst, CSL Behring, Takeda, and Ionis. The rest of the authors declare that they have no no relevant conflicts of interest.

Received for publication May 14, 2022; revised August 28, 2022; accepted for publication August 31, 2022.

Available online November 21, 2022.

Corresponding author: Dawn Laney, MS, CGC, Emory Genetics, 101 Woodruff Circle, Suite 7130, Atlanta, GA 30033. E-mail: Dawn.laney@emory.edu.

2772-8293

https://doi.org/10.1016/j.jacig.2022.08.009

Abbreviations used
EHR: Electronic health record
HAE: Hereditary angioedema
SIS: Suspicion index screening

converting enzyme inhibitors), low C4 complement levels, urticaria/angioedema, cancer, renal disease, cardiac disease, and COVID-19–related edema (n = 33).

A series of 5 logistic regression models with elastic net penalty were fitted for patients with HAE versus negative controls, with consideration for patient groups of C1 inhibitor and patient groups of non–C1 inhibitor that included *PLG*, *KNG1*, *F12*, and *ANGPT1*.[5,6] The elastic net penalty is a weighted combination of L1 (lasso) and L2 (ridge) penalties, with one parameter denoting the proportion of each type of penalty and the other parameter denoting the penalty magnitude.[7,8] We took equal weights for L1 and L2 types of penalties, which has the advantages of enabling variable selection and accounting for highly correlated covariates. The penalty parameter was tuned by 5-fold cross-validation if the number of patients with HAE exceeded 10; otherwise, it was taken as 0.1.[9] The best-tuned penalty parameter was used with all samples to fit a risk prediction model to predict the probability of being an patient with HAE, ranging from 0 to 100 in percentage. Generally, a threshold of 50% can be used to determine whether a test sample is a patient with HAE with risk score higher than 50%. The trained risk prediction models were then applied to independent EHR test data, with 7 patients with HAE (3 C1 inhibitor and 4 not C1 inhibitor) and 7 negative controls.

The C1 inhibitor patient case prediction scoring identified all but 1 patient with C1 inhibitor deficiency and did not identify yield any false positives in the test data. The logistic risk prediction model trained for patients with *PLG* also identified all but 1 of the patients with non–C1 inhibitor and yielded only 1 false-positive result.

To increase risk prediction accuracy and prepare for EHR data sets that may be missing key data points, the number of variables was expanded to 25 (Table I), with each assigned a specific weight using prediction points. The broader variant set does include genetic testing and laboratory values that are diagnostic for HAE to ensure high-specificity values and confirm that laboratory values have resulted in clinical diagnosis and development of appropriate treatment/care plans. The prediction points provide weight to each variable as negative (–100), weak (50), moderate (100), strong (150), or very strong (200) indicators of each genetic condition. Total scores were defined as follows: 200 or more points was defined as an increased risk of having HAE, 100 points defined as possible risk of having HAE, and fewer than 100 defined as no increased risk of having HAE (based on the available data). Individual patient scoring examples of the increased risk versus no increased risk categories, including raw data, are as follows: *increased risk* (total score of 250 points based on available medical records data) is characterized by a low serum C4 level (50 points) plus at least 1 episode of swelling (edema) in the face, lips, tongue, larynx, or throat (100 points) plus an episode of swelling (edema) in the face, lips, tongue, larynx, or throat that does not respond effectively to antihistamines, corticosteroids, or epinephrine (100 points), whereas *no increased risk* (total score of 50 points based on available medical records data) is characterized by swelling (edema) after surgery or a dental procedure (50 points).

After the prediction points were added, the discriminatory power of the SIS tool on the full clinical manifestation database was increased to 100% sensitivity and specificity. The predictive accuracy of the tool was confirmed by using HAE diagnostic laboratory testing guidelines validating that each patient at increased risk met the criteria set out by the US Hereditary Angioedema Association Medical Advisory Board.[1] The threshold of the combined criteria is designed to identify patients as being at high risk if in the clinic, the combination of features would be enough for a clinician familiar with HAE to order diagnostic testing. The tool was calibrated to identify patients requiring further evaluation for HAE and not to diagnose patients.

TABLE I. Final 25-variable set used in the validated SIS tool

Symptoms
● Recurrent swelling in any part of the face, lips, tongue, larynx, extremities, genital, or throat that is nonpruritic
● ≥1 episode of swelling (edema) in face, lips, tongue, larynx, or throat
● Acute swelling episode lasting ≥2 d
● Episode of swelling (edema) in face, lips, tongue, larynx, or throat that does not respond effectively to antihistamines, corticosteroids, or epinephrine
● Recurrent acute abdominal pain and swelling episodes with no fever lasting more than 24 h
● Swelling (edema) after surgery or a dental procedure
● Facial, throat, lip, abdominal edema after starting birth control pills or HRT (estrogen-containing) or pregnancy
● ≥1 emergency room visits for appendicitis
Diagnosis
● Diagnosed with acquired angioedema or ACE inhibitor angioedema
● Diagnosed with HAE/angioedema due to C1 inhibitor disorder/C1 esterase inhibitor deficiency
Family history
● Family history of hereditary angioedema
● Family history of swelling of face, lips, skin, or tongue
● Family member who died of suffocation due to throat or tongue swelling (laryngeal edema)
Genetic testing result
● Pathogenic variant in genes with known association with HAE
○ *ANGPT1* gene mutation
○ *F12* gene (coagulation factor XII) gene mutation
○ *SERPING1* gene mutation
○ *PLG* gene variant
○ *KNG1* gene
Known triggers
● Allergic reaction to ACE inhibitors, causing lip swelling, edema, or abdominal pain
● Allergic reaction during pregnancy that caused lip swelling or edema
Laboratory results
● Low serum C4 level
● Absent or greatly reduced C1 inhibitor level or function
Therapies
● Acute swelling responds to bradykinin B-2 receptor inhibitor or C1 esterase inhibitor
● Treated with HAE-specific medications
● Long-term use of steroids
Negative symptoms
● Hives or urticaria
● Diagnosis of bullous SLE/lupus
● Diagnosis of kidney or heart disease
● Diagnosis of cellulitis

ACE, Angiotensin-converting enzyme; *HRT*, hormone replacement therapy.

RESULTS AND DISCUSSION

The full 25-variable digital SIS tool was run against selected deidentified data fields from 9,981 Ochsner Lafayette General records covering 34,238 encounters over 6 months and 298,288 Emory records covering 12 months of data. At Ochsner Lafayette General, the tool identified 7 patients as being at increased risk for HAE (including angiotensin-converting enzyme inhibitor angioedema) and 297 patients as being at possible risk. The available patient data that scored 100 or higher during this pilot run were reviewed by the study genetic counselors to troubleshoot the tool and determine whether the prediction points of specific variables

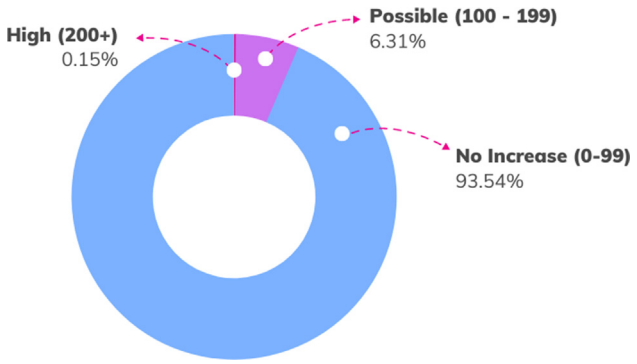

FIG 1. Final distribution of Ochsner Lafayette General data based on total prediction point scores. Categories are as follows: 200 or more points indicates increased risk of having HAE, 100 points indicates possible risk of having HAE, and less than100 points indicates no increased risk of having HAE (based on the available data).

should be adjusted, as well as to evaluate the "real-life" sensitivity and specificity of the system (Fig 1).

At Emory, the tool identified 19 patients as being at increased risk for HAE, including 7 patients with recurrent swelling of the face, lips, tongue, larynx, abdomen, extremities, or throat without hives; 2 patients with a known family history of HAE; 2 individuals treated with HAE-specific medications; and 3 individuals with documented low C1 esterase inhibitor (<21 mg/dL) and low serum C4 levels.

Limitations of the study include the short data review time frame, which could have missed patients not seen during this period of time; the scope of EHR data on a given patient that were available for review; and the inability during this pilot project to provide diagnostic HAE testing for patients to confirm diagnosis.

The discriminatory power of the validated SIS tool on EHR data is strong and can provide an early warning flag that identifies patients at risk for HAE. In practice, an EHR-based clinical message will notify health care providers of the at-risk designation, accompanied by a letter detailing possible next steps for patient evaluation and referral/diagnostic testing options. Partnering with health systems and providers to run the SIS tool on EHR data provides the opportunity for a shortened time to referral, diagnosis, and treatment. These results also suggest that reviewing and documenting family history of angioedema is important in recognizing when a patient could be affected by HAE rather than a multifactorial condition resulting in swelling due to other causes.

Clinical implications: A validated SIS tool run on EHR data can provide an early warning flag that identifies patients at risk for HAE and allows the opportunity for a shortened time to referral, diagnosis, education, and treatment.

REFERENCES

1. Busse PJ, Christiansen SC, Riedl MA, Banerji A, Bernstein JA, Castaldo AJ, et al. US HAEA Medical Advisory Board 2020 guidelines for the management of hereditary angioedema. J Allergy Clin Immunol Pract 2021;9:132-50.e3.
2. Zanichelli A, Longhurst HJ, Maurer M, Bouillet L, Aberer W, Fabien V, et al. Misdiagnosis trends in patients with hereditary angioedema from the real-world clinical setting. Ann Allergy Asthma Immunol 2016;117:394-8.
3. Zanichelli A, Magerl M, Longhurst HJ, Aberer W, Caballero T, Bouillet L, et al. Improvement in diagnostic delays over time in patients with hereditary angioedema: findings from the Icatibant Outcome Survey. Clin Transl Allergy 2018;8:42.
4. Liu S, Wang X, Xu Y, Xu Q, Zhi Y. Risk factors for diagnostic delay in Chinese patients with hereditary angioedema. Allergy Asthma Proc 2019;40:343-9.
5. Hastie T, Tibshirani R, Friedman JS. The elements of statistical learning. 2nd ed. New York: Springer-Verlag; 2009.
6. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol 2005;67:301-20.
7. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Methodol 1996;58:267-88.
8. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 2000;42:80-6.
9. Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B Methodol 1974;36:111-33.