Population-based familial aggregation of eosinophilic esophagitis suggests a genetic contribution

Kristina Allen-Brady, PhD,[a] Rafael Firszt, MD,[b] John C. Fang, MD,[a] Jathine Wong, BS,[c] Ken R. Smith, PhD,[c] and Kathryn A. Peterson, MD[a] *Salt Lake City, Utah*

Background: Prior familial clustering studies have observed an increased risk of eosinophilic esophagitis (EoE) mostly among first-degree relatives, suggesting a genetic contribution to EoE, and twin studies have suggested a powerful contribution from environmental factors.

Objective: This study sought to clarify the contribution of genetic factors to EoE through estimation of familial aggregation and risk of EoE in extended relatives.

Methods: The Utah Population Database, a population-based genealogy resource linked to electronic medical records for health care systems across the state of Utah, was used to identify EoE cases and age, sex, and birthplace-matched controls at a 5:1 ratio. Logistic regression was used to determine the odds of EoE among relatives of EoE probands compared with the odds of EoE among relatives of controls.

Results: There were 4,423 EoE cases and 24,322 controls. The population-attributable risk of EoE was 31% (95% CI, 28% to 34%), suggesting a relatively strong genetic contribution. Risks of EoE were significantly increased among first-degree relatives (odds ratio [OR], 7.19; 95% CI, 5.65-9.14), particularly first-degree relatives of EoE cases diagnosed <18 years of age (OR, 16.3; 95% CI, 9.4-28.3); second-degree relatives (OR, 1.99; 95% CI, 1.49-2.65); and first cousins (OR, 1.35; 95% CI, 1.03-1.77), providing evidence of a genetic contribution. However, spouses of EoE probands were observed to be at increased risk of EoE (OR, 2.86; 95% CI, 1.31-6.25), suggesting either positive assortative mating or a shared environmental contribution to EoE.

Conclusions: This study supports a significant genetic contribution to EoE as evidenced by increased risk of EoE in distant relatives. (J Allergy Clin Immunol 2017;140:1138-43.)

Key words: *Eosinophilic esophagitis, family clustering, familiality analysis, Utah Population Database*

Eosinophilic esophagitis (EoE) is a chronic antigen-driven inflammatory condition of the esophagus that results in diminished quality of life due to dysphagia, chest pain, emesis, abdominal pain, and food bolus impaction. EoE is a newly recognized disease as International Classification of Diseases, Ninth Revision (ICD-9) diagnostic coding for EoE began in 2008. Prior to the 1990s, EoE was rarely recognized and EoE symptoms were thought to be due to reflux esophagitis.[1,2] Pooled prevalence estimates of EoE from 2008 to 2013 after the ICD-9 diagnostic code became available, have been reported to be 26.3 (95% CI, 12.3-45.5) per 100,000 population.[3]

There is evidence that family history is a risk factor for EoE, suggesting a genetic contribution to the disease.[4,5] The risk of EoE in first-degree relatives has been observed to be 10- to 64-fold higher compared with that of the general population, with males being at higher risk than females.[6] However, twins studies have suggested a powerful role for environmental factors. A comparison of the concordance rates for monozygotic twins (57.9 ± 9.5%) and dizygotic twins (36.4 ± 9.3%) was found to be nonsignificant ($P = .11$), and common environment was found to explain a large proportion, estimated at 81%, of variation in the heritability.[6] Clarification of the contribution to EoE by genetic factors is imperative as candidate gene and other genetic studies of EoE are currently underway (see review article by Rothenberg[7]). Familial aggregation as an indicator of genetic contribution can be best examined by study of more distantly related cases (eg, affected cousins) who likely do not share a common environment.

The purpose of this study was to estimate the familial clustering patterns and risk of EoE in both close and distant relatives. We have available a unique, population-based genealogy resource that has been linked to electronic medical records for health care systems across the state of Utah that facilitates investigation of familial clustering by disease status.

METHODS

The Utah Population Database

The Utah Population Database (UPDB)[8] links genealogy information for the state of Utah to inpatient and outpatient electronic health records for the 2 largest health care systems in the state of Utah, which serve approximately 85% of Utah residents, as well as statewide medical data collected by the Utah Department of Health. Genealogy information was obtained from family

From the Departments of [a]Internal Medicine and [b]Pediatrics, and the [c]Huntsman Cancer Institute Utah Population Database Resource, University of Utah.

We acknowledge partial support for the UPDB through grant P30 CA2014 from the Huntsman Cancer Foundation, the University of Utah, and from the University of Utah's Program in Personalized Health and Center for Clinical and Translational Science. UPDB research is also supported in part by the National Center for Research Resources grant, Sharing Statewide Health Data for Genetic Research (R01 RR021746, GM, PI) with additional support from the Utah State Department of Health and the University of Utah.

Disclosure of potential conflict of interest: J. C. Fang serves on the board for Veritract; serves as a consultant for Boston Scientific; and receives grant support from Obalon. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication February 12, 2016; revised November 30, 2016; accepted for publication December 21, 2016.

Available online February 10, 2017.

Corresponding author: Kristina Allen-Brady, PhD, Genetic Epidemiology, 391 Chipeta Way, Suite D, Salt Lake City, UT 84108. E-mail: kristina.allen@utah.edu.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2016.12.979

Abbreviations used
EoE: Eosinophilic esophagitis
ICD-9: International Classification of Diseases, Ninth Revision
OR: Odds ratio
PAC: Probability of causation
PAR: Population-attributable risk
UPDB: Utah Population Database
UUHSC: University of Utah Health Sciences Center Hospital and Clinics

group sheets of Utah pioneers and their descendants that were collected and computerized in the 1970s and combined with birth certificate information (mother, father, and child) to extend the genealogy to the present day. There are over 9 million unique individual records in the UPDB. The majority of families living in Utah are represented in the UPDB, and most families can be linked across at least 5 generations.[8] Studies have shown that the Utah population is mostly of Northern European descent[9] and has experienced high migration rates and low levels of inbreeding among the founding population.[10]

Electronic health records for this study were obtained from the University of Utah Health Sciences Center Hospital and Clinics (UUHSC) and Intermountain Healthcare. Electronic records from UUHSC have been linked to the UPDB back to 1994 for 1.7 million patients. Intermountain Healthcare records have been linked to the UPDB back to 1995 for approximately 4 million records. In addition to electronic medical records from UUHSC and Intermountain Healthcare, statewide hospital discharge summaries (that include comprehensive diagnostic information), collected by the Utah Department of Health and linked to the UPDB, were used for this study. This study was approved by the University of Utah Institutional Review Board and the Resource for Genetic Epidemiological Research, which oversees the UPDB. All data for this project were deidentified, and informed consent from subjects was waived.

Identification of EoE cases

EoE cases were identified using the ICD-9 diagnosis code 530.13 using all available records from 2008 to 2013. Cases were required to be at least 3 years of age at diagnosis of EoE to try to eliminate diagnoses made without endoscopy. As availability of the ICD-9 diagnosis code for EoE is relatively recent, cases with 1 documented diagnosis were included in the analysis. Prior data suggest that the EoE diagnostic code has adequate specificity.[11] As further confirmation of the validity of the ICD-9 code, we reviewed 100 random EoE cases with the ICD-9 diagnostic code of 530.13; EoE was confirmed in 93 of the 100 cases by pathology review (>15 eosinophils per high power field [eos/hpf]) and symptoms of either dysphagia or chest pain refractory to proton pump inhibitors following consensus diagnostic guidelines for EoE.[12] For cases with multiple diagnoses, the first diagnosis date was used to determine age at diagnosis. To be included in this study, subjects were required to have genealogy information for at least 1 parent or 1 child. There were no exclusion criteria for EoE cases.

Population controls

Controls were selected from the UPDB who had family history information available for at least 1 parent or 1 child. They were matched at a ratio of 5:1 to cases by sex, birth year, birth place (Utah or elsewhere), and absence of a known EoE diagnosis (follow-up information is available on the controls at least until the index cases' diagnostic year); controls were selected without replacement (ie, a control could only be used once). Once controls were selected, relatives of the controls (eg, first-degree relatives) were identified based on the UPDB genealogy, and those relatives with an EoE diagnosis were determined. Hence, controls themselves were free of a known diagnosis of EoE, but there was no restriction on the EoE status of their family members.

TABLE I. Demographic characteristics of 4423 EoE probands

Characteristic	Male	Female
Total	2855	1568
Age range at diagnosis (y)	3-92	3-91
Mean (median) age at diagnosis (y)	36.23 (35)	38.55 (37)
Race, n (%)		
White	2543 (89.1)	1406 (89.7)
Nonwhite	60 (2.1)	30 (1.9)
Unknown	252 (8.8)	132 (8.4)
Ethnicity, n (%)		
Hispanic	68 (2.4)	55 (3.5)
Non-Hispanic	2787 (97.6)	1513 (96.5)

Spouses

We also investigated risk of EoE among spouses, who were assumed to be unrelated or distantly related (sixth- or higher degree relative) to the index case or control. Analyses involving spouses were limited to those index cases and controls who had children. Spouses were defined as the married or unmarried coparent with the index cases and controls. Analyses involving sibling spouses also required the sibling and their spouse to have a child, and sibling spouses were defined as spouses of siblings without EoE.

Statistical analysis

An in-house developed software program specifically for analysis of UPDB data as well as the software package R (R version 2.14.2 for Windows 7; R Foundation, Vienna, Austria) was used for all analyses. Familial analyses were done using logistic regression; we report the odds of EoE among relatives of EoE probands compared with the odds of EoE among relatives of the controls. Odds ratios (ORs) are a surrogate for relative risk estimates when the prevalence of the disease is low, as it is for EoE. Model covariates included sex and birth year. All relatives of the EoE cases and matched controls were included systematically in the calculations even if that relative had been counted previously. In families with multiple affected individuals (eg, siblings), each case was included as a separate index case and risk among all siblings of each case was calculated separately. This approach has been shown to lead to an unbiased estimate of familial risk.[13,14] To account for the nonindependence of observations within kindreds and to match appropriate controls to the EoE cases, we used a random-effects model for each kindred group defined by the nearest female ancestor between a pair of relatives (eg, paternal grandmother for paternal cousins). Odds of EoE were investigated in first-degree relatives, which included parents, siblings, and offspring; second-degree relatives, which included grandparents, grandchildren, aunts/uncles, and nieces/nephews; third-degree relatives, which were limited to first cousins; fourth-degree relatives, which were limited to first cousins, once removed; and fifth-degree relatives, which were limited to second cousins. In addition, the odds of EoE in spouses and sibling spouses were estimated. As EoE case information is available only since 2008 and electronic medical records are available since approximately 1995, analyses were focused on intragenerational, horizontal relatives (ie, cousins) rather than vertical relatives (ie, great grandparents) as vertical relatives are less likely to have an EoE diagnosis because of the narrow window of diagnostic data available by calendar year.

The incidence rate of EoE in Utah was calculated by estimating the incidence of EoE by year from 2008 to 2013, dividing it by an estimate of the Utah population over the same time frame, and multiplying it by 100,000. Cumulative prevalence rates of EoE were calculated by summing the number of EoE cases over multiyear time periods, dividing it by the estimated Utah population over the same time periods, and multiplying it by 100,000. The Utah population count was determined using all subjects who linked to the UPDB for a given year and who met the genealogy requirement of having family history information for at least 1 parent or child available in the UPDB. Population-attributable risk (PAR) of EoE, or the proportion of EoE that is related to familial clustering, was calculated using a method by Bruzzi et al.[15] For each EoE proband, the probability that EoE was caused by membership in a kindred was calculated using