

Accepted Manuscript

Population-based Familial Aggregation of Eosinophilic Esophagitis Suggests a Genetic Contribution

Kristina Allen-Brady, PhD, Rafael Firszt, MD, John C. Fang, MD, Jathine Wong, BS, Ken R. Smith, PhD, Kathryn A. Peterson, MD



PII: S0091-6749(17)30227-0

DOI: [10.1016/j.jaci.2016.12.979](https://doi.org/10.1016/j.jaci.2016.12.979)

Reference: YMAI 12636

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 12 February 2016

Revised Date: 30 November 2016

Accepted Date: 21 December 2016

Please cite this article as: Allen-Brady K, Firszt R, Fang JC, Wong J, Smith KR, Peterson KA, Population-based Familial Aggregation of Eosinophilic Esophagitis Suggests a Genetic Contribution, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2016.12.979.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Population-based Familial Aggregation of Eosinophilic Esophagitis Suggests a Genetic Contribution

Kristina Allen-Brady PhD¹, Rafael Firszt MD², John C. Fang MD¹, Jathine Wong BS³,

Ken R. Smith PhD³, Kathryn A. Peterson MD¹

Departments of ¹Internal Medicine and ²Pediatrics,
and the ³Huntsman Cancer Institute Utah Population Database Resource
University of Utah, Salt Lake City, Utah, 84108 USA

Financial Support: Internal Funds

Corresponding author:

Kristina Allen-Brady
Genetic Epidemiology
391 Chipeta Way, Suite D
Salt Lake City, UT 84108
Phone: 801-585-5333
Fax: 801-581-6052
E-mail: kristina.allen@utah.edu

Abstract

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: 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 healthcare 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 to 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 28-34%), suggesting a relatively strong genetic contribution. Risks of EoE were significantly increased among first-degree relatives (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.

Conclusion: This study supports a significant genetic contribution to EoE as evidenced by increased risk of EoE in distant relatives.

46 **Keywords:** eosinophilic esophagitis, family clustering, familiarity analysis, Utah Population

47 Database

48 Key Messages

- 49 • Excess familial clustering of eosinophilic esophagitis (EoE) observed out to first cousins
50 (third degree relative) provides evidence of a genetic contribution to the disease.
- 51 • First-degree relatives of EoE cases diagnosed at an early age (<18 years) are at greatly
52 increased risk of EoE.
- 53 • A small, but statistically significant number of spouses of EoE probands were affected
54 with EoE providing evidence of either positive assortative mating or a shared
55 environmental contribution to EoE.
- 56

57 **Capsule summary:** This study provides evidence of a genetic contribution to eosinophilic
58 esophagitis as indicated by excess familial clustering of the disease in both close and distant
59 relatives.

Introduction

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 ICD-9 diagnostic coding for EoE began in 2008. Prior to the 1990's 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 to the general population, with males being at higher risk than females.⁶ However, twin studies have suggested a powerful role for environmental factors. A comparison of the concordance rates for monozygotic twins (57.9% $\pm$ 9.5%) and dizygotic twins (36.4% $\pm$ 9.3%) was found to be non-significant ($p=0.11$), and common environment was found to explain a large proportion of variation in the heritability, estimated at 81%.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 article⁷). Familial aggregation as an indicator of genetic contribution can be best examined by study of more distantly related cases (e.g., 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 healthcare systems across the state of Utah which facilitates investigation of familial clustering by disease status.

Materials and Methods

The Utah Population Database (UPDB)

The UPDB⁸ links genealogy information for the state of Utah to inpatient and outpatient electronic health records for the two largest healthcare 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 group sheets of Utah pioneers and their descendants collected and computerized in the 1970's 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. Majority of families living in Utah are represented in the UPDB, and most families can be linked across at least five generations.⁸ Studies have shown that the Utah population is mostly of Northern European descent⁹ 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 (IH). 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 IH, 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 de-identified 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 three 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 one documented diagnosis were included in the analysis. Prior data suggests 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 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 one parent or one 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 one parent or one 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 (i.e., a control could only be used once). Once controls were selected, relatives of the controls (e.g., 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.

Spouses

We also investigated risk of EoE among spouses, who were assumed to be unrelated or distantly related (6th 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 co-parent 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) was used for all analyses. Familial analyses were done using logistic regression; we report the odds of EoE among relatives of EoE probands compared to the odds of EoE among relatives of the controls. Odds ratios 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 (e.g., 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 non-independence 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 (e.g., paternal grandmother for paternal cousins). Odds of EoE were investigated in first-degree relatives (FDR) which included parents, siblings, and offspring; second-degree relatives (SDR) 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 (i.e., cousins) rather than vertical relatives (i.e., 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 timeframe, and multiplying it by 100,000. Cumulative prevalence rates of EoE were calculated by summing the number of EoE cases over multi-year 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 one 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 the probability of causation (PAC) formula $PAC = (OR-1)/OR$, where the OR is determined for each individual as a function of familial standardized incidence ratio. The PAR is calculated as the mean of the PAC across all index cases.

High-risk EoE pedigrees were determined by identifying founders with an excess number of affected descendants compared with the expected number of affected descendants based on population rates of EoE based on age- and sex-matched cohorts ($p < 0.01$). High-risk pedigrees were further determined by requiring that the pedigree's familial standardized incidence ratio be greater than 5, which accounts for the number of relatives in the pedigree, degree of relatedness to the proband, and person-time at risk among family members.¹⁶

Results

There were 5,138 unique EoE subjects identified in the UPDB, 4,423 of these subjects had at least one relative in the UPDB and were defined as probands for this study. These 4,423 EoE cases were age, sex, and birthplace-matched to 24,322 controls. Demographic characteristics of cases by sex are displayed in Table 1. Nearly twice as many diagnosed EoE cases were male. The mean age at diagnosis was 36.2 years for males and 38.6 years for females.

Incidence rates and cumulative prevalence rate by year and by age (<18 and ≥ 18 years) are shown in Table 2. Relatively few EoE cases were diagnosed in 2008, which is not unexpected given the newness of the ICD-9 diagnostic code at that time. The largest increase in the number of incident cases occurred in 2009. The cumulative population prevalence of EoE in

Utah by 2013 was 28.45 cases per 100,000 population, which is similar to a pooled published prevalence rate of EoE since 2008 of 26.3 (95% CI 12.3 - 45.5) per 100,000 population.³

For the 4,423 EoE probands, there were 230 observed affected first-degree relatives (1.17%) compared to 165 EoE cases among matched controls (0.18%) for an adjusted odds ratio of 7.19 (95% CI 5.65 – 9.14, $p=3.8e-58$, Table 3). Table 3 also shows the number of unaffected relatives for the EoE cases and matched controls; for example there were 19,455 unaffected first-degree relatives of 4,423 EoE probands and 91,517 unaffected first-degree relatives of 24,322 matched controls. A significantly increased risk of EoE was observed in second-degree relatives (OR 1.99, 95% CI 1.49-2.65, $p=2.5e-06$), and first cousins (OR 1.35, 95% CI 1.03-1.77, $p=0.03$), providing evidence of a genetic contribution for EoE as familial clustering extended to even distant relatives. An example of a high-risk EoE pedigree is shown in Figure 1. Significant risk of EoE did not extend to first cousins once removed ($p=0.94$) or second cousins ($p=0.31$). There was an increased risk of EoE among spouses of EoE probands (OR 2.86, 95% CI 1.31-6.25, $p=0.0086$), but not among spouses of EoE siblings ($p=0.094$), suggesting either positive assortative mating or a contribution from shared environmental factors to EoE. However, the total number of spouses (10 total individuals or 5 husband-wife pairs) who were both affected with EoE was small (0.39%).

Further exploration of the risk of EoE in relatives stratified by age of the proband (<18 and ≥ 18) is shown in Table 4. There were 813 probands diagnosed with EoE prior to the age of 18 years identified. Among these earlier diagnosed EoE probands, there was a greatly increased risk of EoE among first-degree relatives of any age (OR 16.3, 95% CI 9.4-28.3, $p=3.3e-23$) and an increased risk of EoE among second-degree relatives (OR 2.02, 95% CI 1.06-3.86, $p=0.032$), but no increased risk was observed for other more distant relatives. When first-degree relatives

of earlier diagnosed probands were restricted to also be less than 18 years of age (data not shown in Table), the odds of having a diagnosis of EoE increased to 34.44 (95% CI 15.5-76.54, $p < 2 \times 10^{-16}$) suggesting that siblings of younger EoE cases are at greatly increased risk of EoE. For probands greater than or equal to 18 years of age at diagnosis of EoE ($n=3,610$), an increased risk of EoE for first-degree relatives, second-degree relatives, and first cousins was observed. Inclusion of an interaction term in the logistic regression model between age of proband (<18 versus ≥ 18) and affection status of first-degree relative (case versus control) was found to be significant ($p=2.25 \times 10^{-4}$), suggesting a significant difference in risk of EoE to first-degree relatives of probands diagnosed at an earlier age.

As males are more likely to be affected with EoE, the odds of diagnosis of EoE among male relatives compared to female relatives in our Utah population was explored (Table 5). As can be seen, males have approximately 1.9 times increased risk of being diagnosed with EoE compared to females across all relative types and controls.

The population attributable risk for familial factors contributing to EoE was estimated to be 31% (95% CI 28% - 34%). The population attributable risk represents the contribution from familial or genetic factors.

Discussion

In this Utah population-based study, we observed evidence of a significant familial clustering of EoE in first-degree relatives, second-degree relatives, and first cousins (a third-degree relative). While familial EoE cases may be attributable to shared genetic factors or shared

environmental factors, the excess clustering of EoE in distant relatives, who are less likely to share immediate environments, suggests a genetic contribution to the disease.

A number of prior studies investigating the family history risk of EoE mostly among first-degree relatives have observed an increased risk of EoE.^{6, 17-21} Our study is the first to explore risk of EoE in distant relatives. We report the population attributable risk for familial factors, which indicates the proportion of EoE related to familial causes, to be 31%, which represents a relatively strong contribution from familial/genetic factors. For comparison, the PAR for EoE is similar to the PAR for familial factors for colon cancer calculated using the UPDB, the Utah Cancer Registry, and the same methodology as was used for this study at 32% (95% CI 26 - 37%).²² However, given that a substantial proportion of the PAR for EoE is not explained by familial factors, our results are also consistent with a strong contribution from environmental factors.⁶

Study of high-risk families with an excess of EoE may facilitate EoE predisposition gene identification. A number of candidate gene studies and genome-wide association studies are being conducted to identify EoE risk variants.⁷ Extended families with multiple disease cases are a proven means for finding disease-causing genes.²³ Given that environmental effects contribute to EoE, cases recruited from high-risk families, such as the example high-risk pedigree shown in Figure 1, may reduce the chance of sampling sporadic clusters and increase the frequency of shared rare variants within the family and hence increase power to detect the causal variant(s).²⁴

Early onset disease is sometimes indicative of stronger genetic effects (e.g., cancer or Alzheimer's disease).^{25, 26} However, in this study for early onset probands we did not observe

stronger genetic effects as evidenced by an increased risk of EoE in both close and distant relatives. We observed a greatly increased risk of EoE among first-degree relatives of probands with earlier diagnosed disease (age <18 years), a risk of EoE in second-degree relatives that was comparable between probands diagnosed at age <18 years and age $\geq$ 18 years, and no increased risk of EoE in first cousins. This increased risk of EoE in first-degree relatives of earlier onset probands is most likely because siblings with similar symptoms could be more easily diagnosed through prior knowledge of the disease by their parents and treating physicians as well as referrals to specialists who treat EoE, though work remains to validate this explanation. The lack of an increased risk of EoE in more distant relatives may be due to multiple factors including small sample size of earlier onset cases or a relatively new diagnostic code for EoE; further investigation of the risk of EoE in distant relatives of earlier diagnosed probands in the future is warranted as awareness of EoE increases over time.

The increased risk of EoE observed in this study among a small number of unrelated, adult, spouse pairs (5 total spouse pairs, or 10 subjects) is noteworthy. The number of years of marriage prior to EoE diagnosis for these five marriages ranged between 6 years and 36 years, and all five of the marriages occurred prior to 2008. Given that the ICD-9 code for EoE started in 2008 and that individuals in this study could not have received the ICD-9 diagnostic code prior to 2008, we cannot determine whether the increased risk of EoE observed for spouses is due to shared environmental contributions or assortative mating. Future research involving affected spouse pairs may help to understand this observed increased risk for EoE.

Similar to other studies,^{27, 28} we also observed an increased risk of EoE in males. One prior study reported a cohort of EoE patients in which 75% were male.²⁹ A nonsynonymous single nucleotide polymorphism associated with the thymic stromal lymphopoietin (TSLP)

receptor gene has been found to be involved in the pathogenesis of EoE; this gene is encoded in the pseudoautosomal regions on the X and Y-chromosomes and may account, in part, for the increased frequency in males.³⁰ It is also possible that there are sex-specific epigenetic modifications that influence gene expression levels resulting in male predominance of EoE.

There are limitations of this study. We used a single occurrence of an ICD-9 code indicating EoE to identify cases. While EoE is a relatively new diagnosis and prior studies indicate that the EoE diagnostic code has adequate specificity,¹¹ it is possible that some misclassification bias may be present in this study. As a check on misclassification bias, the crude EoE prevalence in controls was reported (Table 3), which should approximate EoE prevalence in the general population. In this study, the crude EoE prevalence in controls ranged from 0.18% to 0.23%. In a study by Ronkainen et al., which estimated prevalence of EoE in a community-based study in Sweden,³¹ upper endoscopy was performed on 1,000 randomly selected adults; using an EoE diagnostic criterion of > 20 eos/hpf, they found prevalence of definite EoE to be 0.4%. The prevalence of EoE in this Swedish community is very similar to our crude EoE prevalence among controls. Given that our sample size of EoE cases was substantial and population-based, the comparability of control prevalence rates suggests that we accurately quantified EoE disease risk in relatives. Another limitation of this study is that since EoE is a relatively new diagnostic code, there may be many undiagnosed EoE cases within the various healthcare systems used for this study possibly due to delays in diagnosis of EoE, misdiagnosis, as well as symptomatic subjects without a diagnosis of EoE who have not sought medical attention (e.g., invasive diagnostic endoscopy required). However, despite the possibility of missing individuals with EoE, our study prevalence estimate is similar to other published prevalence estimates,³ providing empiric validity for our study. We assumed that

missed EoE diagnostic rates were similar between relatives of cases and relatives of controls since they were drawn from the same sample; however, close relatives of an EoE case may be more likely to have a diagnosis of EoE because of prior knowledge of EoE within the family.

The findings from this study are generalizable to other populations of Western and Northern European ancestry. Studies using the UPDB genealogy resource have shown that the Utah population is mostly of northern European descent⁹ and has experienced high migration rates and low levels of inbreeding among the founding population.³²

In conclusion, we have identified a significant genetic contribution to EoE that extends to distant relatives. Study of families with an excess of EoE may facilitate EoE predisposition gene identification and hence further understanding of EoE etiology.

Acknowledgements

We thank the Pedigree and Population Resource of the Huntsman Cancer Institute, and the University of Utah (funded in part by the Huntsman Cancer Foundation) for its role in the ongoing collection, maintenance and support of the Utah Population Database (UPDB). We thank the University of Utah Center for Clinical and Translational Science (CCTS) (funded by NIH Clinical and Translational Science Awards), the Pedigree and Population Resource, University of Utah Information Technology Services and Biomedical Informatics Core for establishing the Master Subject Index between the Utah Population Database, the University of Utah Health Sciences Center and Intermountain Health Care. We also 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 CCTS. UPDB research is also supported in part by the NCRR grant, "Sharing Statewide Health Data for Genetic Research" (R01 RR021746, G. Mineau, PI) with additional support from the Utah State Department of Health and the University of Utah

329 **List of Abbreviations**

330 EoE: eosinophilic esophagitis

331 95% CI: 95% confidence interval

332 ICD-9: International classification of Disease version 9

333 IH: Intermountain Healthcare

334 OR: odds ratio

335 PAC: probability of causation

336 PAR: population attributable risk

337 UPDB: Utah Population Database

338 UUHSC: University of Utah Health Sciences Center Hospital and Clinics

339

340

References

1. Attwood SE, Smyrk TC, Demeester TR, Jones JB. Esophageal eosinophilia with dysphagia. A distinct clinicopathologic syndrome. *Dig Dis Sci* 1993; 38:109-16.
2. Kelly KJ, Lazenby AJ, Rowe PC, Yardley JH, Perman JA, Sampson HA. Eosinophilic esophagitis attributed to gastroesophageal reflux: improvement with an amino acid-based formula. *Gastroenterology* 1995; 109:1503-12.
3. Arias A, Perez-Martinez I, Tenias JM, Lucendo AJ. Systematic review with meta-analysis: the incidence and prevalence of eosinophilic oesophagitis in children and adults in population-based studies. *Aliment Pharmacol Ther* 2016; 43:3-15.
4. Noel RJ, Putnam PE, Rothenberg ME. Eosinophilic esophagitis. *N Engl J Med* 2004; 351:940-1.
5. Blanchard C, Wang N, Rothenberg ME. Eosinophilic esophagitis: pathogenesis, genetics, and therapy. *J Allergy Clin Immunol* 2006; 118:1054-9.
6. Alexander ES, Martin LJ, Collins MH, Kottyan LC, Sucharew H, He H, et al. Twin and family studies reveal strong environmental and weaker genetic cues explaining heritability of eosinophilic esophagitis. *J Allergy Clin Immunol* 2014; 134:1084-92 e1.
7. Rothenberg ME. Molecular, genetic, and cellular bases for treating eosinophilic esophagitis. *Gastroenterology* 2015; 148:1143-57.
8. Utah Population Database. Available from <http://www.huntsmanccancer.org/research/shared-resources/utah-population-database/overview>.
9. McLellan T, Jorde LB, Skolnick MH. Genetic distances between the Utah Mormons and related populations. *Am J Hum Genet* 1984; 36:836-57.

10. Jorde LB, Skolnick M. Demographic and genetic applications of computerized record linkage: the Utah Mormon genealogy. *Informatique et Sciences Humaines* 1981; 56:105-17.
11. Rybnicek DA, Hathorn KE, Pfaff ER, Bulsiewicz WJ, Shaheen NJ, Dellon ES. Administrative coding is specific, but not sensitive, for identifying eosinophilic esophagitis. *Dis Esophagus* 2014; 27:703-8.
12. Liacouras CA, Furuta GT, Hirano I, Atkins D, Attwood SE, Bonis PA, et al. Eosinophilic esophagitis: updated consensus recommendations for children and adults. *J Allergy Clin Immunol* 2011; 128:3-20 e6; quiz 1-2.
13. Bai Y, Sherman S, Khoury MJ, Flanders WD. Bias associated with study protocols in epidemiologic studies of disease familial aggregation. *Am J Epidemiol* 2000; 151:927-37.
14. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics* 2000; 56:645-6.
15. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk factors using case-control data. *Am J Epidemiol* 1985; 122:904-14.
16. Kerber RA. Method for calculating risk associated with family history of a disease. *Genet Epidemiol* 1995; 12:291-301.
17. Collins MH, Blanchard C, Abonia JP, Kirby C, Akers R, Wang N, et al. Clinical, pathologic, and molecular characterization of familial eosinophilic esophagitis compared with sporadic cases. *Clin Gastroenterol Hepatol* 2008; 6:621-9.
18. Katzka DA. Eosinophilic esophagitis: it's all in the family. *Gastrointest Endosc* 2007; 65:335-6.

19. Meyer GW. Eosinophilic esophagitis in a father and a daughter. *Gastrointest Endosc* 2005; 61:932.
20. Patel SM, Falchuk KR. Three brothers with dysphagia caused by eosinophilic esophagitis. *Gastrointest Endosc* 2005; 61:165-7.
21. Zink DA, Amin M, Gebara S, Desai TK. Familial dysphagia and eosinophilia. *Gastrointest Endosc* 2007; 65:330-4.
22. Kerber RA, O'Brien E. A cohort study of cancer risk in relation to family histories of cancer in the Utah population database. *Cancer* 2005; 103:1906-15.
23. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009; 461:747-53.
24. Blangero J, Diego VP, Dyer TD, Almeida M, Peralta J, Kent JW, Jr., et al. A kernel of truth: statistical advances in polygenic variance component models for complex human pedigrees. *Adv Genet* 2013; 81:1-31.
25. Alberici A, Benussi A, Premi E, Borroni B, Padovani A. Clinical, genetic, and neuroimaging features of Early Onset Alzheimer Disease: the challenges of diagnosis and treatment. *Curr Alzheimer Res* 2014; 11:909-17.
26. Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. *Gut* 2015; 64:1623-36.
27. Kapel RC, Miller JK, Torres C, Aksoy S, Lash R, Katzka DA. Eosinophilic esophagitis: a prevalent disease in the United States that affects all age groups. *Gastroenterology* 2008; 134:1316-21.

28. Prasad GA, Talley NJ, Romero Y, Arora AS, Kryzer LA, Smyrk TC, et al. Prevalence and predictive factors of eosinophilic esophagitis in patients presenting with dysphagia: a prospective study. *Am J Gastroenterol* 2007; 102:2627-32.
29. Spergel JM, Brown-Whitehorn TF, Beausoleil JL, Franciosi J, Shuker M, Verma R, et al. 14 years of eosinophilic esophagitis: clinical features and prognosis. *J Pediatr Gastroenterol Nutr* 2009; 48:30-6.
30. Sherrill JD, Gao PS, Stucke EM, Blanchard C, Collins MH, Putnam PE, et al. Variants of thymic stromal lymphopoietin and its receptor associate with eosinophilic esophagitis. *J Allergy Clin Immunol* 2010; 126:160-5 e3.
31. Ronkainen J, Talley NJ, Aro P, Storskrubb T, Johansson SE, Lind T, et al. Prevalence of oesophageal eosinophils and eosinophilic oesophagitis in adults: the population-based Kalixanda study. *Gut* 2007; 56:615-20.
32. Jorde LB. Inbreeding in the Utah Mormons: an evaluation of estimates based on pedigrees, isonymy, and migration matrices. *Ann Hum Genet* 1989; 53:339-55.

425 Table 1. Demographic characteristics of 4,423 EoE probands

Characteristic	Male	Female
Total	2,855	1,568
Age range at diagnosis	3-92 years	3-91 years
Mean (median) age at diagnosis	36.23 (35) years	38.55 (37) years
Race: white (%)	2,543 (89.1%)	1,406 (89.7%),
non-white (%)	60 (2.1%)	30 (1.9%)
unknown (%)	252 (8.8%)	132 (8.4%)
Ethnicity: Hispanic (%), non-Hispanic (%)	68 (2.4%), 2,787 (97.6%)	55 (3.5%), 1,513 (96.5%)

426

427

Table 2. Incidence and cumulative prevalence rates per 100,000 population of EoE by year and age.

Year	< 18 years (n=813)			≥18 years (n=3,610)			Cumulative Population Prevalence
	EoE cases	UPDB Population	Incidence Rate	EoE cases	UPDB Population	Incidence Rate	
2008	41	646,108	6.35	165	1,856,378	8.89	8.23
2009	129	672,209	19.19	585	1,858,546	31.48	18.28
2010	109	677,534	16.09	676	1,843,892	36.66	22.57
2011	108	689,288	15.67	814	1,920,082	42.39	25.85
2012	213	698,146	30.51	670	1,964,002	34.11	27.37
2013	213	730,026	29.18	700	1,992,234	35.14	28.45

Table 3. Risk of EoE in relatives of 4,423 EoE probands and 24,322 age, sex, and birthplace-matched controls

Relative type	Relatives of Cases		Relatives of controls		OR (95% CI)	p-value
	Affected / Unaffected Relatives	% relatives affected	Affected / Unaffected Relatives	% relatives affected		
First-degree	230 / 19,455	1.17%	165 / 91,517	0.18%	7.19 (5.65–9.14)	3.8e-58
Second-degree	100 / 33,111	0.30%	251 / 152,663	0.16%	1.99 (1.49–2.65)	2.5e-06
First cousins	109 / 43,747	0.25%	324 / 196,859	0.16%	1.35 (1.03–1.77)	0.03
First cousins, once removed	316 / 151,656	0.21%	1142 / 680,819	0.17%	0.99 (0.85–1.16)	0.94
Second cousins	510 / 256,535	0.2%	2152 / 1,166,308	0.18%	0.94 (0.85–1.05)	0.31
Spouses	10 / 2,546	0.39%	17 / 12,218	0.14%	2.86 (1.31–6.25)	0.0086
Sibling Spouses	16 / 10,980	0.15%	124 / 52,888	0.23%	0.60 (0.33–1.06)	0.094

Table 4. Risk of EoE in relatives stratified by age of proband

Age of Proband	Relative type	Relatives of Cases		Relatives of controls		OR (95% CI)	p-value
		Affected / unaffected relatives	% relatives affected	Affected / unaffected relatives	% relatives affected		
<18 yrs n=813							
	First-degree	77 / 2,829	2.65%	25 / 13,627	0.18%	16.3 (9.4-28.3)	3.3e-23
	Second-degree	16 / 4,505	0.35%	52 / 20,143	0.26%	2.02 (1.06-3.86)	0.032
	First cousins	8 / 5,913	0.14%	39 / 26,650	0.15%	0.86 (0.33-2.24)	0.75
	First cousins once removed	46 / 16,833	0.27%	152 / 72,578	0.21%	1.08 (0.67-1.73)	0.75
	Second cousins	56 / 27,459	0.2%	176 / 116,500	0.15%	1.17 (0.77-1.77)	0.47
≥18 yrs n=3,610							
	First-degree	153 / 16,626	0.91%	140 / 77,890	0.18%	5.61 (4.25-7.40)	2.7e-34
	Second-degree	84 / 28,606	0.29%	199 / 132,520	0.15%	2.07 (1.50-2.84)	7.7e-06
	First cousins	101 / 37,834	0.27%	285 / 170,494	0.17%	1.43 (1.07-1.90)	0.015
	First cousins once removed	270 / 134,823	0.20%	990 / 608,241	0.16%	1.04 (0.88 – 1.23)	0.66
	Second cousins	454 / 229,076	0.20%	1,976 / 1,049,808	0.19%	0.94 (0.83 – 1.05)	0.27

439 Table 5. Risk of EoE in male relatives compared to female relatives

Relative type	Relatives of Cases		Relatives of controls		OR (95% CI)	p-value
	Affected / Unaffected Male Relatives	% male relatives affected	Affected / Unaffected Male Relatives	% male relatives affected		
First-degree	135 / 9,723	1.37%	120 / 45,366	0.26%	1.96 (1.53-2.50)	2.55e-09
Second-degree	66 / 16,489	0.40%	158 / 75,410	0.21%	1.94 (1.52-2.48)	1.33e-07
First cousins	76 / 22,138	0.34%	213 / 99,668	0.21%	1.98 (1.58-2.48)	4.05e-09
First cousins once removed	225 / 76,023	0.30%	779 / 340,570	0.23%	2.20 (1.92-2.52)	1.47e-29
Second cousins	349 / 128,905	0.27%	1,435 / 587,830	0.24%	1.99 (1.82-2.18)	2.38e-50
Spouse	5 / 893	0.56%	9 / 4,576	0.20%	1.60 (0.74-3.43)	0.23
Sibling spouse	8 / 5,509	0.15%	80 / 26,639	0.30%	1.96 (1.31-2.94)	0.0011

440

441

Figure Legends

Figure 1. Example of a high-risk EoE pedigree. The founders of the pedigree have 1,207 descendants; 6 were diagnosed with EoE, whereas only 1.03 were expected based on age, sex, and birthplace matched population rates ($p=0.0007$). Males are represented as squares, females as circles, and shaded diamonds represent cases. The diamond shape, which is a unisex designation, was used to protect the identity of the pedigree. As a high-risk pedigree, the number of affected individuals significantly exceeds the number of cases expected.

