J ALLERGY CLIN IMMUNOL
JANUARY 2021

Polygenic risk score for atopic dermatitis in the Canadian population

To the Editor:

Atopic dermatitis (AD) is characterized by a damaged skin barrier that allows allergens to penetrate the body, leading to sensitization and a higher risk of developing food allergies (relative risk [RR], 33.79), asthma (RR, 7.04), and/or rhinitis (RR, 11.75), all features of the atopic march.[1] Recent evidence has shown that the atopic march can be modified in high-risk infants with early interventions directed at reestablishing and/or maintaining skin barrier function with intense use of simple emollients, and introducing food allergens early into the diet.[2-5] Although these constitute examples of low-intensity, high-impact interventions for health care systems, their successful and indiscriminate implementation in the whole population is neither feasible nor realistic. In this context, building a predictive tool to identify children at high risk of developing moderate to severe AD (MSAD) would allow targeted interventions with maximized impact. In this study, a polygenic risk score (PRS) with an area under the curve (AUC) of 88% and explaining 37% of MSAD variance was established for the Canadian population.

Two scenarios for PRS were tested, one using genome-wide association study (GWAS) loci identified through existing literature and the other based on the strongest GWAS hits found in 2 Canadian cohorts (see Table E1 in this article's Online Repository at www.jacionline.org; for the detailed methodology, see this article's Methods section in the Online Repository at www.jacionline.org).

The first scenario evaluated whether the best associations in the literature were suitable to build a PRS for AD with a good discriminative value in a specific population. The 25 best associations documented in GWASs of AD (see Tables E2 and E3 in this article's Online Repository at www.jacionline.org) were selected. For each of these, a region spanning ±100 kb was tested for the best associated genetic variant with MSAD (for clinical definitions, see this article's Online Repository at

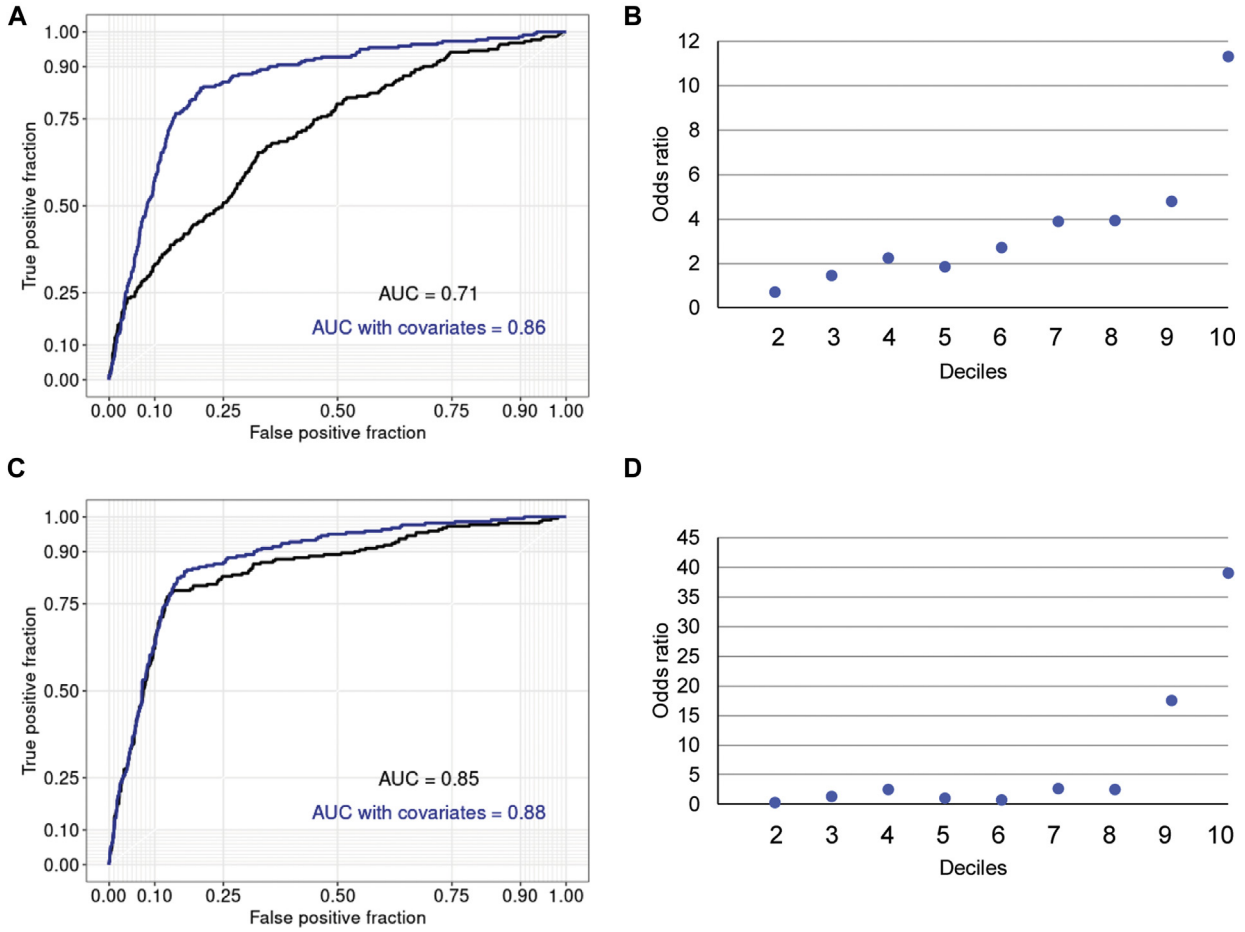

FIG 1. Receiver-operating characteristic (ROC) curves and AUCs for the PRSs of MSAD, calculated for associated single nucleotide polymorphisms in the literature (**A**), and in 2 Canadian cohorts (**C**), along with the comparison between deciles 2 to 10 and decile 1 for both PRSs in (**B**) and (**D**). Dark lines represent the ROC curves for models without the covariates and blue lines for those with them.

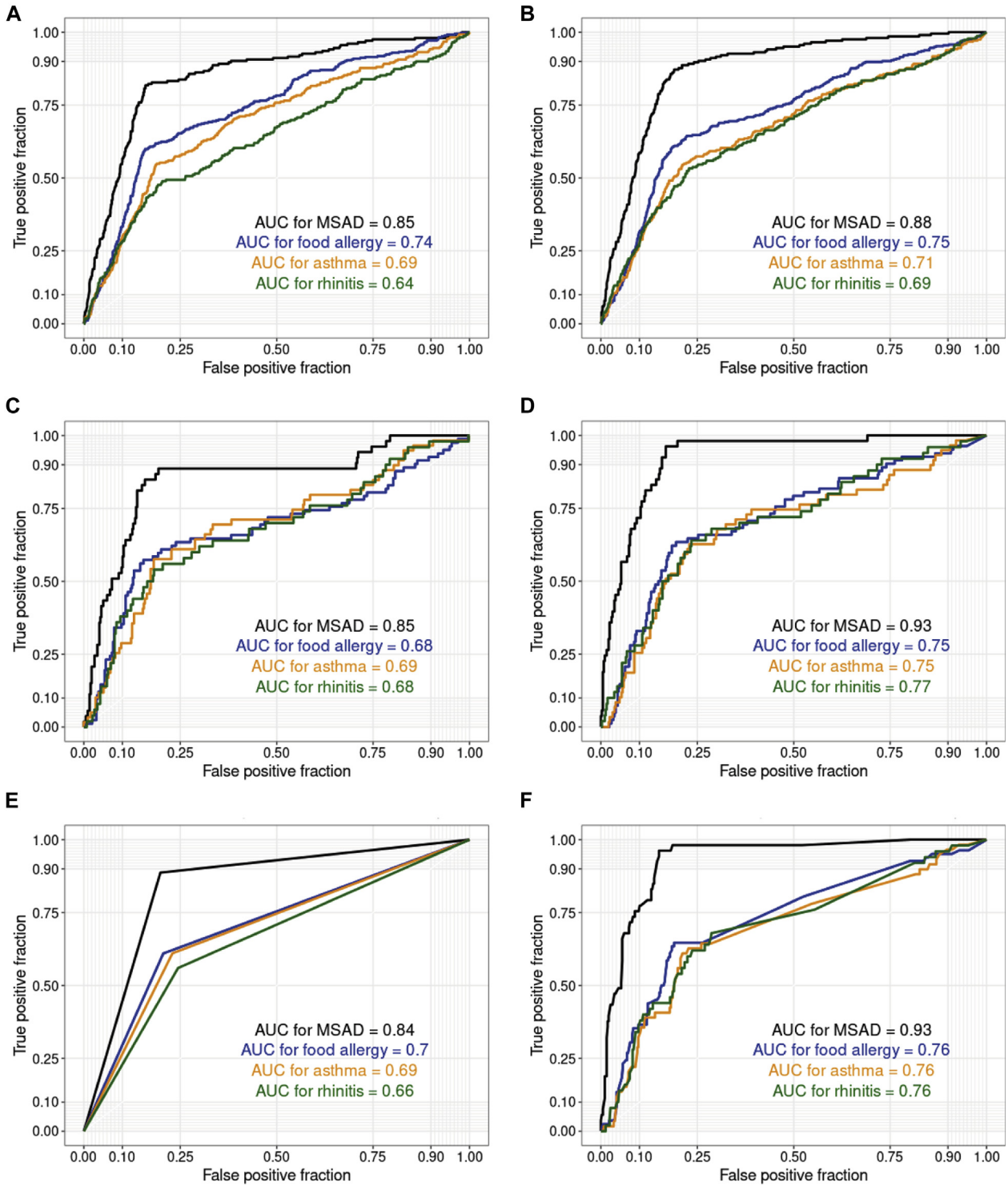

FIG 2. Receiver-operating characteristic (ROC) curves and AUCs for diseases in the atopic march: MSAD (black), food allergies (blue), asthma (orange), and rhinitis (green). The PRS was calculated from the 25 best genome-wide associations in the training cohort for the model without (**A**) and with (**B**) covariates, and in the testing cohort for the model without (**C**) and with (**D**) covariates. Finally, a model using the binary variable derived from the PRS was run in the testing cohort for the model without (**E**) and with (**F**) covariates.

www.jacionline.org) in 80% of the unrelated cases and controls from the 2 Canadian cohorts (training cohort, n = 2688 individuals; see Fig E1 in this article's Online Repository at www.jacionline.org) using a general regression model to extract risk alleles and β estimates (= ln[odds ratio]).[E7] The PRS was then built for each individual of the training and testing groups (the remaining 20% of individuals; n = 676) considering the number of risk alleles from genetic variants weighted by their β estimates. The discriminative value of PRS was assessed by a receiver-operating characteristic curve analysis and gave an AUC of 71% (Fig 1, *A*). According to Nagelkerke's pseudo-R^2, 11% of MSAD variance was explained by PRS. Once covariates were added to the model (sex, age, and parents' ethnicity), an AUC of 86% was reached (Fig 1, *A*) and the model explained 31% of MSAD variance. These results highlight the dependence of the first scenario upon covariates to reach a good discriminative value (AUC between 80% and 89%).[E12]

The second scenario took advantage of 2 Canadian cohorts to build another PRS based on the best associations for AD in the Canadian population, which is ethnically diverse. The Saguenay–Lac-Saint-Jean asthma familial cohort[E1] includes individuals of French descent from this region in northern Quebec, Canada, and the CHILD cohort study[E2] comprises both those of English descent and those of multiple other origins living in British Columbia (Vancouver), Alberta (Edmonton), Manitoba (Winnipeg, Morden, and Winkler), and Ontario (Toronto). Each site obtained local Research Ethics Board approval for the study, and each participating parent gave signed informed consent. The GWAS was performed on the training cohort using the Multiple Family-based Quasi-Likelihood Score (MFQLS) test,[E10] and a PRS was built with the 25 best associations (see Table E4 and Fig E1 in this article's Online Repository at www.jacionline.org) after running a general regression model to extract their β estimates. The PRS obtained had an AUC of 85% (Fig 1, *C*) and explained 33% of MSAD variance. Adding the same covariates as for the first scenario, the AUC was 88% (Fig 1, *C*), corresponding to a sensitivity and specificity of both 84%, and the PRS explained 37% of MSAD variance. The 2 models of this second scenario demonstrate a good discriminative value.[E12] In comparison, AUCs derived from PRSs in the literature range from 53% to 76%,[6] and show relatively low percentages of explained variance (0.3% for brain tumor compared with 36% in this study).[7] The PRS based on the Canadian cohorts also carried good predictive values for other allergic phenotypes (for clinical definitions, see this article's Online Repository at www.jacionline.org), with AUCs of 75% for food allergies, 71% for asthma, and 69% for allergic rhinitis (Fig 2, *A* and *B*). Interestingly, the discriminative value for asthma was similar than that reported using the Predicting Asthma Risk in Children clinical tool, based on respiratory symptoms occurring before school age (AUC = 77%), even though the predictive tool proposed in this study is designed to identify children at high risk of developing MSAD and not directly to detect those at high risk of developing asthma.[8] Results were validated in the testing cohort, with AUCs of 85% and 93% for the models for MSAD without and with covariates, respectively (with 31% and 49% of explained variability) and AUCs of 75%, 75%, and 77% for food allergies, asthma, and rhinitis (Fig 2, *C* and *D*).

These results demonstrate that the second scenario, which used data from the targeted population to build the PRS, best explains the risk of developing MSAD even without considering any covariate. It is interesting to note that no locus was common between the 25 best associations from the 2 scenarios (Tables E2-E4). It confirms the need to characterize the genetic profile of each specific population before building a PRS in order to reach a good discriminative value.

To be an efficient predictive tool, a cutoff value has to be established to distinguish between low- and high-risk individuals. When examining the progression of risk to develop MSAD on the basis of individuals' PRSs from the second scenario, there is a clear inflection in the curve at the 9th decile (OR, 17.5, compared with the first decile) with further progressing up to the 10th decile (OR, 39.0; Fig 1, *D*; see Table E5 in this article's Online Repository at www.jacionline.org). In contrast, risk progression was much more continuous with PRSs calculated from the first scenario, which is further evidence of its lower discriminative value (Fig 1, *B*, and Table E5). Using a binomial logistic regression with the above-mentioned covariates in the training cohort, being classified as high risk (defined as having a PRS between the 9th and 10th deciles) was strongly associated with the risk of developing AD (odds ratio [OR], 2.96; $P = 2.62 \times 10^{-07}$), MSAD (OR, 11.73; $P = 1.27 \times 10^{-21}$), food allergies (OR, 5.80; $P = 2.96 \times 10^{-19}$), asthma (OR, 3.96; $P = 2.12 \times 10^{-10}$), and rhinitis (OR, 2.29; $P = .001$; see Table E6 in this article's Online Repository at www.jacionline.org). For comparison, a PRS built from 4 selected genes (*GSTP1*, *TNF*, *TLR2*, and *TLR4*) was previously reported to have an OR of only 1.22 for AD.[9] Analyses in the testing cohort also gave significant results for MSAD (OR, 5.67; $P = .001$), food allergies (OR, 3.27; $P = .003$), and asthma (OR, 3.15; $P = .013$) even though the cohort used was smaller. Moreover, AUCs for the binary PRS in the testing cohort had similar values than for the continuous PRS (Fig 2, *E* and *F*). These are interesting results for clinical applications because such a predictive tool based on only 25 genetic variants and basic covariates available at birth could be easily added to the test routines already established in Canada, which are performed on blood samples collected by a prick on the heels of newborns.

Finally, to test whether a smaller number of genetic variants can be as efficient, a PRS was tested for genetic variants from the second scenario that were associated with P less than 1×10^{-10} (n = 8). The discriminative value was well preserved with a corresponding AUC of 86% (Fig E2). However, analyses using a cutoff value to identify children at high risk of developing MSAD in the testing cohort gave less interesting results than the ones for the PRS built from 25 genetic variants, showing a greater dependence on covariates (AUC for the model without covariates = 67% and with covariates = 92%).

To conclude, the use of 2 independent multiethnic Canadian cohorts allowed development of a PRS with an AUC of 88% and explaining 37% of MSAD variance. Considering the accumulating body of evidence indicating the need to intervene early to prevent the development of AD and associated allergic comorbidities, discriminative PRS such as this one could prove helpful to guide interventions and direct investments toward those patients most likely to benefit, ensuring the cost-effectiveness and sustainability of early prevention programs.

We thank the participants recruited in the Saguenay–Lac-Saint-Jean asthma cohort and the CHILD cohort for their valuable participation in this study. We also thank Dominique Fournier, scientific language expert, for the revision of this manuscript (http://www.serviceslinguistiquesdf.com/home).

Mathieu Simard, BSc[a,b]
Anne-Marie Madore, PhD[a]
Simon Girard, PhD[a,b]
Susan Waserman, MD, MSc[c]
Qingling Duan, PhD[d]
Padmaja Subbarao, MD, MSc[e]
Malcolm R. Sears, MBChB[f]
Theo J. Moraes, MD, PhD[e]
Allan B. Becker, MD[g]
Stuart E. Turvey, MBBS, DPhil[h]
Piushkumar J. Mandhane, MD, PhD[i]
Charles Morin, MD[a,j]
Philippe Bégin, MD, PhD[k,l]
Catherine Laprise, PhD[a,b,j]

From [a]Département des sciences fondamentales and [b]Centre intersectoriel en santé durable, Université du Québec à Chicoutimi, Saguenay, Quebec, [c]the Division of Clinical Immunology and Allergy, McMaster University, Hamilton, [d]the School of Computing and Department of Biomedical & Molecular Sciences, Queen's University, Kingston, [e]the Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, and [f]the Division of Respirology, McMaster University, Hamilton, Ontario, [g]the Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Manitoba, [h]the Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, [i]the Department of Pediatrics, University of Alberta, Edmonton, Alberta, and [j]the Department of Pediatrics, Centre intégré universitaire de santé et de services sociaux du Saguenay–Lac-Saint-Jean, Saguenay, [k]the Department of Medicine, Centre hospitalier de l'Université de Montréal, Montreal, and [l]the Department of Pediatrics, CHU Sainte-Justine, Montreal, Quebec, Canada. E-mail: catherine.laprise@uqac.ca.

M.S. is supported by a *Fonds de recherche du Québec - Santé* (FRQS) Master's Training Award. This study had support from the Quebec Respiratory Health Network (RHN; https://rsr-qc.ca/en/) pilot project grants and by the Canada Research Chair in the Environment and Genetics of Respiratory Disorders and Allergy. The Canadian Institutes of Health Research (CIHR), and the Allergy, Genes and Environment (AllerGen) Network of Centres of Excellence provided core support for the CHILD study. C.L. is part of the Quebec RHN, the investigator of the CHILD study, the director of the *Centre intersectoriel en santé durable de l'UQAC* and the chairholder of the Canada Research Chair in the Environment and Genetics of Respiratory Disorders and Allergy (http://www.chairs.gc.ca). The GWAS data were made available by the European Commission as part of GABRIEL (A multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract number 018996 under the Integrated Program LSH-2004-1.2.5-1 (Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control). This study makes use of data generated by the UK10K Consortium. A full list of the investigators who contributed to the generation of the data is available from www.UK10K.org. Funding for UK10K was provided by the Wellcome Trust under award WT091310.

Disclosure of potential conflict of interest: The authors declare that they have no relevant conflicts of interest.

REFERENCES

1. Tran MM, Lefebvre DL, Dharma C, Dai D, Lou WYW, Subbarao P, et al. Predicting the atopic march: results from the Canadian Healthy Infant Longitudinal Development Study. J Allergy Clin Immunol 2018;141:601-7.e8.
2. Du Toit G, Sayre PH, Roberts G, Sever ML, Lawson K, Bahnson HT, et al. Effect of avoidance on peanut allergy after early peanut consumption. N Engl J Med 2016; 374:1435-43.
3. Horimukai K, Morita K, Narita M, Kondo M, Kitazawa H, Nozaki M, et al. Application of moisturizer to neonates prevents development of atopic dermatitis. J Allergy Clin Immunol 2014;134:824-30.e6.
4. Natsume O, Kabashima S, Nakazato J, Yamamoto-Hanada K, Narita M, Kondo M, et al. Two-step egg introduction for prevention of egg allergy in high-risk infants with eczema (PETIT): a randomised, double-blind, placebo-controlled trial. Lancet 2017;389:276-86.
5. Simpson EL, Chalmers JR, Hanifin JM, Thomas KS, Cork MJ, McLean WH, et al. Emollient enhancement of the skin barrier from birth offers effective atopic dermatitis prevention. J Allergy Clin Immunol 2014;134:818-23.
6. Do CB, Hinds DA, Francke U, Eriksson N. Comparison of family history and SNPs for predicting risk of complex disease. PLoS Genet 2012;8:e1002973.
7. Adel Fahmideh M, Lavebratt C, Tettamanti G, Schuz J, Roosli M, Kjaerheim K, et al. A weighted genetic risk score of adult glioma susceptibility loci associated with pediatric brain tumor risk. Sci Rep 2019;9:18142.
8. Pedersen ESL, Spycher BD, de Jong CCM, Halbeisen F, Ramette A, Gaillard EA, et al. The simple 10-item Predicting Asthma Risk in Children Tool to predict childhood asthma—an external validation. J Allergy Clin Immunol Pract 2019;7: 943-53.e4.
9. Huls A, Klumper C, MacIntyre EA, Brauer M, Melen E, Bauer M, et al. Atopic dermatitis: interaction between genetic variants of GSTP1, TNF, TLR2, and TLR4 and air pollution in early life. Pediatr Allergy Immunol 2018; 29:596-605.

Available online May 19, 2020.
https://doi.org/10.1016/j.jaci.2020.04.057

Eosinophilic esophagitis with extremely high esophageal eosinophil counts

To the Editor:

Eosinophilic esophagitis (EoE) is an inflammatory disease characterized by eosinophil infiltration into the esophageal mucosa, with a peak count of at least 15 eosinophils per hpf following endoscopic biopsy. However, the range of esophageal eosinophilia can vary markedly from patient to patient. A key question in the field is to understand the relationship of eosinophil levels with disease features, especially because eosinophil-targeted therapies are now available. Patients with extremely high levels of esophageal eosinophilia have not previously been studied. It is unknown whether these patients exhibit characteristics different from those of patients with EoE who have esophageal eosinophilia that is near the threshold of disease diagnosis. Given this fact, we aimed to establish whether any significant clinical, endoscopic, histologic, or transcriptomic features differ between patients with extremely high levels of esophageal eosinophilia and those with levels near the threshold of disease diagnosis.

Among those in a registry of patients with EoE, we identified the group of patients with the highest recorded levels of esophageal eosinophilia (>350 eosinophils per hpf), referred to as EoE-High. We subsequently identified a second group that had relatively low levels of esophageal eosinophilia (15-24 eosinophils per hpf), referred to as EoE-Low. There were 74 patients in the registry with eosinophil counts of 15 to 24 eosinophils per hpf on a distal esophageal biopsy specimen. A random number generator was used to select the 14 patients comprising the EoE-Low group. Phenotypic and clinical characteristics were gathered on the basis of electronic medical records and detailed questionnaires as part of a research registry. Endoscopic characteristics were assessed on the basis of findings from esophagogastroduodenoscopy operative reports. Histologic characteristics were classified on the basis of the histology scoring system.[1] Molecular analysis was performed by using the 96-gene EoE diagnostic panel.[2] Age, demographics, esophageal eosinophil levels, absolute eosinophilia, IgE levels, atopic comorbidities, and treatment modalities were assessed. Study population characteristics are summarized in Table E1 (in this article's Online Repository at www.jacionline.org). At the time of biopsy, the EoE-Low group was younger than the EoE-High group, with mean ages of 6.3 ± 3.4 years and 13.4 ± 10.3 years, respectively ($P = .02$). The EoE-High group had a disease duration significantly longer than that in the