Association of Environmental Factors with Age-Related Macular Degeneration using the Intelligent Research in Sight Registry

Matthew S. Hunt, BS,[1,3] *Yewlin E. Chee, MD,*[1,2] *Steven S. Saraf, MD,*[1,2] *Emily Y. Chew, MD,*[4]
Cecilia S. Lee, MD, MS,[1,2] *Aaron Y. Lee, MD, MSCI,*[1,2] *Michael B. Manookin, PhD*[1,2]

Purpose: Investigate associations of natural environmental exposures with exudative and nonexudative age-related macular degeneration (AMD) across the United States.

Design: Database study.

Participants: Patients aged ≥ 55 years who were active in the IRIS Registry from 2016 to 2018 were analyzed. Patients were categorized as nonexudative, inactive exudative, and active exudative AMD by International Classification of Diseases 10th Revision and Current Procedural Terminology (CPT) codes. Patients without provider-level ZIP codes matching any ZIP code tabulation area were excluded.

Methods: Environmental data were obtained from public sources including the US Geological Survey, National Renewable Energy Laboratory, National Oceanic and Atmospheric Administration, and Environmental Protection Agency. Multiple variable, mixed effects logistic regression models with random intercepts per ZIP code tabulation area quantified the association of each environmental variable with any AMD versus non-AMD patients, any exudative AMD versus nonexudative AMD, and active exudative AMD versus inactive exudative and nonexudative AMD using 3 separate models, while adjusting for age, sex, race, insurance type, smoking history, and phakic status.

Main Outcome Measure: Odds ratios for environmental factors.

Results: A total of 9 884 527 patients were included. Elevation, latitude, solar irradiance measured in global horizontal irradiance (GHI) and direct normal irradiance (DNI), temperature and precipitation variables, and pollution variables were included in our models. Statistically significant associations with active exudative AMD were GHI (odds ratio [OR], 3.848; 95% confidence interval [CI] with Bonferroni correction, 1.316−11.250), DNI (OR, 0.581; 95% CI, 0.370−0.913), latitude (OR, 1.110; 95% CI, 1.046−1.178), ozone (OR, 1.014; 95% CI, 1.004−1.025), and nitrogen dioxide (OR, 1.005; 95% CI, 1.000−1.010). The only significant environmental associations with any AMD were inches of snow in the winter (OR, 1.005; 95% CI, 1.001−1.009) and ozone (OR, 1.011; 95% CI, 1.003−1.019).

Conclusions: The strongest environmental associations differed between AMD subgroups. The solar variables GHI, DNI, and latitude were significantly associated with active exudative AMD. Two pollutant variables, ozone and nitrogen dioxide, also showed positive associations with AMD. Further studies are warranted to investigate the clinical relevance of these associations. Our curated environmental dataset has been made publicly available at https://github.com/uw-biomedical-ml/AMD_environmental_dataset. *Ophthalmology Science 2022;2:100195 © 2022 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.ophthalmologyscience.org.

Age-related macular degeneration (AMD) is a common neurodegenerative disease of the retina and a major cause of vision loss in people aged 50 years and older.[1,2] Although we currently lack a detailed understanding of the complete etiology and pathogenesis of AMD, both genetic and environmental factors likely contribute to the development and progression of the disease.[3−9] Thus, a major goal of AMD research is to identify environmental factors that contribute to disease progression and use this knowledge to develop effective prevention strategies.

Environmental variables such as smoking have shown consistent associations with the prevalence of AMD,[10−12] but it is unclear whether other natural factors such as sunlight exposure contribute to the development and progression of this disease. For example, the studies on association between light exposure and AMD have shown conflicting results. Their limited sample size or diversity in population may have contributed to these discrepancies.[13,14] Thus, large population-based datasets in combination with reliable environmental information are needed to determine

whether environmental variables such as light exposure are associated with increased risk of development or progression of AMD.

The Intelligent Research in Sight (IRIS®) Registry provides large-scale, population-level data across the United States, capturing disease diagnoses, patient demographic information, and multiple metrics of ocular health. Environmental information was extracted from publicly available government sources, including the US Geological Survey, National Renewable Energy Laboratory, National Oceanic and Atmospheric Administration, and Environmental Protection Agency. These sources provide information about several natural environmental factors across the United States, including elevation, solar irradiance, precipitation, and pollution. Coupling the patient-level information from the IRIS Registry with the environmental data provided a unique opportunity to investigate whether environmental factors contribute to AMD.

Methods

Methods for collecting data from the IRIS registry have been previously described.[15] This study was exempted from the University of Washington Institutional Review Board because of de-identified data use. The version of the IRIS database used for this study (Rome 2.0) was last modified on October 23, 2020.

Patient Selection

The study population included all patients aged 55 years or older who were actively being followed in the IRIS registry at any point from January 1, 2016, to December 31, 2018. Patients with AMD were identified using International Classification of Diseases 9th or 10th Revision codes and grouped into 3 mutually exclusive outcomes. Nonexudative AMD patients were identified by the presence of at least 1 AMD code, but no exudative AMD codes from 2016 to 2018. Patients with active exudative AMD had at least 1 active exudative International Classification of Diseases code from 2016 to 2018 and received anti−VEGF (injections at any time from 2016 to 2018, as identified by at least 1 Current Procedural Terminology (CPT) code for intravitreal injection, 67028). Patients with inactive exudative AMD received at least 1 exudative AMD code, but did not have a CPT code for intravitreal injection or did not have an active exudative AMD code. Patient location information was determined on the basis of the postal code of the first provider who had postal code data. Patients were excluded if their providers did not have postal code information or if their provider's postal code did not match a valid ZIP Code Tabulation Area (ZCTA). Patient-level variables included patient age, race, gender, insurance type, smoking history, and phakic status. Patient insurance categories included commercial, Medicare, Medicare Advantage, unknown insurance, no insurance, and other insurance. Patients with any prior history of smoking were identified as those having former or active "[t]obacco use and exposure…" in the IRIS Registry. Patients were determined to be pseudo/aphakic if they received a CPT code for cataract surgery or an International Classification of Diseases code for pseudo/aphakia before the beginning of 2016.

ZCTA and Environmental Variables

ZIP Code Tabulation Areas are representations of the areas of ZIP codes.[16] The US Census Bureau calculates the latitude and longitude of ZCTA "internal points," which are approximations of the geographic center of ZCTAs.[17] ZIP Code Tabulation Areas and internal points were obtained from the Tigris package in R.[18] The internal point of each ZCTA was used to define a single latitude and longitude location for each ZCTA. Each ZCTA internal point was matched to the nearest location of measurement for each environmental variable to assign each patient within each ZCTA a single set of environmental measurements. Matching of internal points to environmental coordinates was performed using Haversine distance, or distance between 2 points on a sphere, with Sci-Kit Learn's BallTree algorithm.[19]

All environmental factor data were obtained from publicly available government sources. Elevation information was obtained from the US Geological Survey's "3DEP LidarExplorer."[20] These data were a 1-arcsecond resolution Digital Elevation Model dataset, which consisted of tiles seamlessly covering the contiguous United States.[21] This dataset contained minimum bounding boxes for each region and the mean elevation of each region, in meters. The upper-left and lower-right latitude and longitude coordinates of the minimum bounding box were used to determine a centroid of each Digital Elevation Model region. These centroids were used to map each ZCTA to an elevation.

Solar irradiance data in the form of global horizontal irradiance (GHI) and direct normal irradiance (DNI) were obtained from the National Renewable Energy Laboratory's National Solar Radiation Database.[22] Global horizontal irradiance represents the irradiance incident upon a horizontal (relative to the ground) collecting surface. Direct normal irradiance represents the irradiance directly coming from the solar disc, incident upon a collecting surface always held normal to the sun.[23] Direct normal irradiance thus excludes scatter irradiance from the atmosphere or other sources of scatter. Illustrations of GHI and DNI are shown in Figure S1 (available at www.ophthalmologyscience.org). The GHI and DNI values were calculated based on the multi-year Physical Solar Model developed by NASA, which takes into account cloud properties and water vapor. The resolution of the GHI and DNI data was 0.038 degrees in both latitude and longitude ($\sim 4 \times 4$ km).[24]

Weather data were obtained from the National Oceanic and Atmospheric Administration's 1981−2010 Climate Normals.[25] Weather data included 12 different measurements each averaged over the 30 years from 1981 to 2010. Six of these measurements were for temperature throughout the year, and 6 were for precipitation throughout the year. The variables annual precipitation and summer precipitation were combined into a single "other precipitation" variable, and the same was done for the annual and summer snow variables. The coordinates of the 4807 weather stations that had complete measurements for all temperature and precipitation variables were used to map each ZCTA to a set of weather measurements.

Pollution data were obtained from the Environmental Protection Agency's 2017 Annual Summary Data.[26] Of the many pollutants recorded in that data, 7 pollutants (carbon monoxide, lead, ozone, PM2.5, PM10, nitrogen dioxide, and sulfur dioxide) were selected for analysis based on their relevance to National Ambient Air Quality Standards.[27] The latitude and longitude of stations recording levels of these pollutants were included in the data and used to map each ZCTA to a set of pollution measurements. Of these pollutants, lead was excluded from the analysis due to a paucity of measurement locations, and PM10 was excluded because it strongly correlated with PM2.5 on exploratory data analysis. The final pollutants analyzed thus included carbon monoxide, ozone, PM2.5, nitrogen dioxide, and sulfur dioxide. All environmental factors used in modeling are also shown

plotted across the contiguous United States in Figure S2 (available at www.ophthalmologyscience.org). To enable future analyses with these environmental factors, we have open-sourced our dataset after combining these public resources at https://github.com/uw-biomedical-ml/AMD_environmental_dataset.

Statistical Models

Three separate mixed-effects logistic regression models were used to quantify the association of environmental variables with each of our 3 outcomes: any AMD, any exudative AMD, and active exudative AMD. The cohort populations differed between the models and are illustrated in Figure S3 (available at www.ophthalmologyscience.org). The binary outcomes of the 3 models were any AMD versus no AMD, which included all patients meeting inclusion criteria, any exudative AMD versus nonexudative AMD, and active exudative AMD versus inactive-exudative AMD and nonexudative AMD. The latter 2 models included all AMD patients. Each model used all patient-level variables, environmental variables, and latitude as predictors. A logistic regression model was fit as a Generalized Additive Model using the BAM package in R[28], treating postal code as a random effect using a random intercept per postal code to account for cluster effects by ZIP code. The Bonferroni correction was used to alter the confidence interval (CI) size for model parameters to adjust for multiple comparisons (98.3% CI used).

Results

We studied the contribution of environmental factors to AMD by analyzing environmental and other factors for approximately 10 million patients listed in the IRIS registry. The total number of patients tracked in the IRIS registry from January 1, 2016, to December 31, 2018, with either year of birth less than or equal to their date of activity in the registry or null year of birth was 51 983 471, of whom 48 394 167 had no AMD, 2 613 559 had non-exudative AMD, 389 403 had inactive exudative AMD, and 586 342 had active exudative AMD. After applying all inclusion criteria, the total number of patients with no AMD, nonexudative AMD, inactive exudative AMD, and active exudative AMD were 8 599 964, 913 552, 124 562, and 246 449, respectively. Figure 1 shows a patient-inclusion flow diagram.

Of the patients included in the final analysis without AMD, the majority were female (58%), and the most common race was White (69%). The mean age was 68 years, and interquartile range (IQR) was 62−74. Pseudo/aphakic patients (21%) were less common than phakic patients, and patients with a history of smoking (37%) were less common than nonsmokers. Compared with the study population, the patients with nonexudative AMD, inactive exudative AMD, and active exudative AMD had a higher proportion of women (60%, 58%, and 59%, respectively), Whites (80%, 83%, and 87%, respectively), pseudo/aphakic patients (33%, 37%, and 37%, respectively), and smokers (43%, 50%, and 55%, respectively), and tended to be older (mean age 75 years, IQR, 69−80; 77 years, IQR, 71−81 years; and 77 years, IQR, 72−81, respectively). Table 1 shows details on patient demographics for our study population.

Associations among baseline demographic variables, insurance type, smoking history, and phakic status with our outcome are shown in Figure S4 (available at www.ophthalmologyscience.org). Patient-level factors with statistically significant positive association across all models
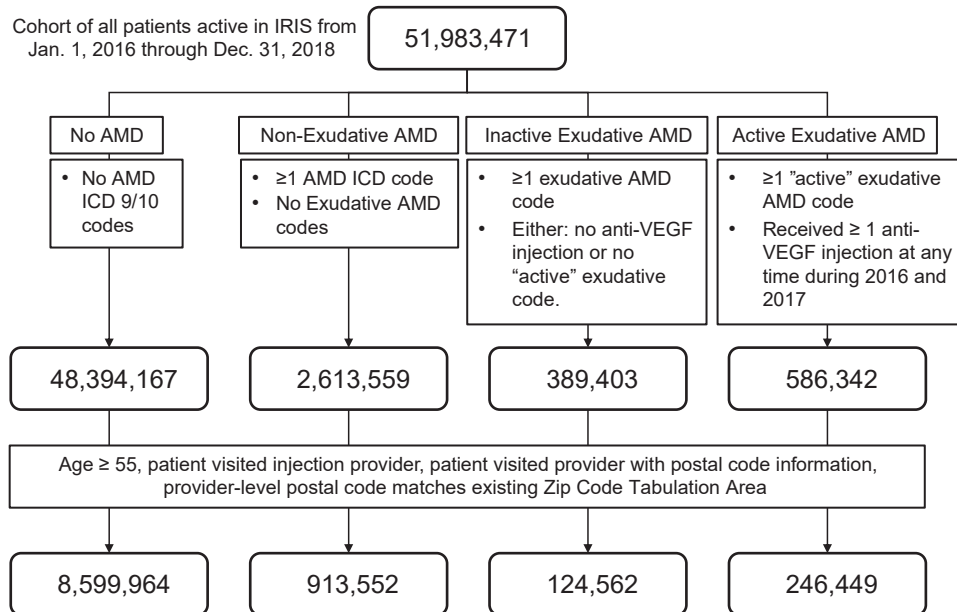

Figure 1. Patient inclusion flowchart. All unique patients active in the Intelligent Research in Sight (IRIS) Registry (51 983 501) during 2016 or 2017 are split into their outcome groups to give the initial raw number of patients in each category. After applying all inclusion criteria, the final numbers are shown. AMD = age-related macular degeneration; ICD 9/10 = International Classification of Diseases 9th Revision and 10th Revision.

Table 1. Demographics

Characteristic	Overall N = 9 884 527	No AMD N = 8 599 964	Nonexudative AMD N = 913 552	Inactive Exudative AMD N = 124 562	Active Exudative AMD N = 246 449	P Value
Age, yrs	69 (63−75)	68 (62−74)	75 (69−80)	77 (71−81)	77 (72−81)	< 0.001
Sex						< 0.001
Female	5 726 680 (58%)	4 963 046 (58%)	545 804 (60%)	72 650 (58%)	145 180 (59%)	
Male	4 123 644 (42%)	3 608 654 (42%)	363 719 (40%)	51 238 (41%)	100 033 (41%)	
Not Reported	34 203 (0.3%)	28 264 (0.3%)	4029 (0.4%)	674 (0.5%)	1236 (0.5%)	
Race						< 0.001
White	6 994 423 (71%)	5 943 743 (69%)	732 652 (80%)	103 196 (83%)	214 832 (87%)	
Black or African American	681 946 (6.9%)	648 117 (7.5%)	27 086 (3.0%)	3151 (2.5%)	3592 (1.5%)	
Asian	242 215 (2.5%)	214 238 (2.5%)	21 497 (2.4%)	2727 (2.2%)	3753 (1.5%)	
Other	94 496 (1.0%)	83 230 (1.0%)	7661 (0.8%)	1353 (1.1%)	2252 (0.9%)	
Unknown	1 871 447 (19%)	1 710 636 (20%)	124 656 (14%)	14 135 (11%)	22 020 (8.9%)	
Insurance Type						< 0.001
Commercial	5 731 197 (58%)	5 038 976 (59%)	497 606 (54%)	64 739 (52%)	129 876 (53%)	
Medicare	1 698 954 (17%)	1 417 570 (16%)	197 167 (22%)	29 876 (24%)	54 341 (22%)	
Medicare Advantage	1 226 480 (12%)	1 020 197 (12%)	144 178 (16%)	19 122 (15%)	42 983 (17%)	
Other	339 266 (3.4%)	298 790 (3.5%)	24 556 (2.7%)	4438 (3.6%)	11 482 (4.7%)	
Medicaid	187 310 (1.9%)	171 916 (2.0%)	10 977 (1.2%)	1616 (1.3%)	2801 (1.1%)	
No Insurance	22 546 (0.2%)	20 615 (0.2%)	1185 (0.1%)	215 (0.2%)	531 (0.2%)	
Unknown	678 774 (6.9%)	631 900 (7.3%)	37 883 (4.1%)	4556 (3.7%)	4435 (1.8%)	
Smoker	3 760 846 (38%)	3 168 728 (37%)	392 833 (43%)	62 858 (50%)	136 427 (55%)	< 0.001
Pseudo/Aphakic	2 199 366 (22%)	1 765 800 (21%)	297 261 (33%)	46 049 (37%)	90 256 (37%)	< 0.001

AMD = age-related macular degeneration.
Raw numbers and percentages in parentheses. Percentages are relative to the total patients in each category (top row) of table.
Median (IQR); n (%).
Kruskal−Wallis rank-sum test; Pearson's chi-square test.

included age, positive smoking history, pseudo/aphakic status, and having "Other" insurance (which included government, military, and miscellaneous insurance types). Statistically significant protective factors included non-White race and male sex.

After adjusting for multiple comparisons, demographic variables, insurance type, smoking history, phakic status, and the random effects of each postal code, the environmental factors with strongest associations differed among the 3 outcomes. Inches of snow in the winter and ozone in parts per billion were the only statistically significant associations in the modeling of any AMD versus non-AMD patients (winter inches of snow OR, 1.005; Bonferroni-corrected 95% CI, 1.001−1.009), ozone (OR, 1.011; 95% CI, 1.003−1.019). For the model of active exudative AMD versus inactive exudative and nonexudative AMD, the statistically significant associations were GHI (OR, 3.848; 95% CI with Bonferroni correction, 1.316−11.250), DNI (OR, 0.581; 95% CI, 0.370−0.913), latitude (OR, 1.110; 95% CI, 1.046−1.178), ozone (OR, 1.014; 95% CI, 1.004−1.025), and nitrogen dioxide (OR, 1.005; 95% CI, 1.000−1.010). Modeling any exudative AMD versus nonexudative AMD did not exhibit any statistically significant environmental associations. The strength of association between the environmental variables and our outcomes is shown in Figure 2.

Predicted risk ratio for development of active exudative AMD was mapped at a high resolution across the United States for a 75-year-old, pseudophakic, White woman with no smoking history and Medicare insurance compared with the same patient with mean environmental values. Risk was predicted at each GPS location and displayed using a Voronoi diagram across the contiguous United States (Fig 3). To reveal which components of our model gave the US risk map this appearance, risk was plotted across the contiguous US using subsets of environmental variables with the other variables held constant. These modified risk-prediction maps are shown in Figure S5 (available at www.ophthalmologyscience.org).

Discussion

In this work, we investigated associations of natural environmental factors and pollution with the development and progression of AMD. Our analysis confirmed previous reports that demographic factors such as sex and ethnicity are strongly correlated with AMD, because many of these factors were significantly correlated in all 3 models. Our analysis further revealed significant associations between several environmental variables and the risk of active exudative AMD, including latitude, solar irradiance, and pollution. However, the environmental associations differed between the subtypes of AMD, with the risk of developing any AMD versus no AMD significantly associated only with inches of winter snow and ozone, and the risk of developing any exudative AMD versus all nonexudative AMD not significantly associated with any environmental factors. Our findings reveal previously unappreciated associations between environmental factors and disease progression in AMD.
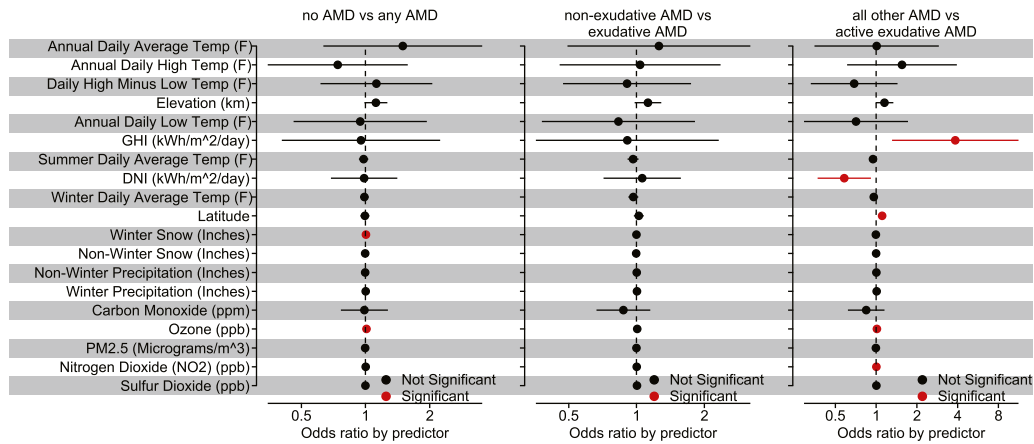
Figure 2. Odds ratios and Bonferroni-corrected 95% confidence intervals (CIs) for predictor variables for each multivariable, mixed-effects, logistic regression model. **Dashed vertical lines**: null hypothesis values; red: statistically significant predictors (global horizontal irradiance [GHI], direct normal irradiance [DNI], latitude, ozone, and nitrogen dioxide for "All Other AMD vs Active Exudative AMD" model). AMD = age-related macular degeneration.

Differences in AMD Subtypes

The development of exudative AMD from nonexudative AMD represents a clear progression in disease severity, and thus motivates the distinction between exudative and non-exudative forms of AMD. The motivation to distinguish active exudative AMD from inactive exudative AMD is not as obvious, but the distinction is important to make given the data available in the IRIS Registry. Approximately 32% of all patients with any exudative AMD (cohort before inclusion criteria Fig 1) did not receive intravitreal injection therapy from 2016 to 2018. Therefore, we split the "any exudative AMD" outcome group into inactive and active exudative AMD to avoid contamination of treatment-requiring exudative AMD with either former cases of exudative AMD now quiescent or end-stage scarring no longer warranting treatment, or erroneous diagnoses of exudative AMD. The requirements for both at least 1 intravitreal injection CPT code (67028) and at least 1 active-exudative International Classification of Diseases code from 2016 to 2018 were used to purify our "active exudative AMD" outcome group to ensure we captured the exudative, advanced form of the disease actively requiring injections.

Precise definitions of outcome groups allowed us to analyze differences in environmental associations along the spectrum of AMD progression and provided a possible explanation for differences between prior work and our study. Previous studies presented conflicting results for associations between solar irradiance and AMD risk, with some data supporting an association between light exposure and AMD,[13] and other data indicating that this association is not significant.[14] The disparity in these findings may partly arise from the studies' limited sample sizes. However, our results indicate that the relationship between solar irradiance and the risk of developing exudative AMD is more nuanced than previously thought. The 2014 study by Delcourt et al[13] found an association between UV exposure and early AMD, but they had only a small number of cases of "late" AMD, which included exudative AMD as well as geographic atrophy, and they

did not find any association between UV exposure and late AMD development.[13] We found measures of solar irradiance (discussed in detail below), along with elevation and latitude were associated with "active exudative AMD" but were not associated with "any AMD" or "any exudative AMD" (Fig 2). This discrepancy between AMD subtypes may indicate a difference in the etiologies of different stages of AMD and could account for the difference in findings between our study and the prior work by Delcourt et al.[13]

Solar Irradiance and AMD risk

We also found that specific types of solar irradiance are associated with AMD in nuanced ways. Separating solar irradiance into GHI (direct + scatter irradiance incident on a horizontal collecting surface) versus DNI (direct-only irradiance incident on a collecting surface perpendicular to incident rays) revealed that these distinct classes of sunlight had opposing associations with the risk of developing active exudative AMD. Direct sunlight (DNI) decreased the risk of developing active exudative AMD, whereas indirect plus direct sunlight (GHI) increased the risk (Fig 2). The retrospective nature of this study precludes causal inference, and the mechanism of the opposing associations between GHI and DNI cannot be determined by this study. However, we consider several speculative mechanisms, based on light wavelength, ocular versus skin exposure, and behavioral modification in response to light.

One hypothesis for the divergent effects of GHI and DNI on AMD risk involves solar angle. Global horizontal irradiance can be calculated from DNI and diffuse horizontal irradiance (DHI), the irradiance on a horizontal collecting surface not coming from the solar disc (scatter irradiance). Global horizontal irradiance is given by DNI times cosine of the solar angle away from vertical, theta, plus the DHI.[29] Ignoring the DHI component, GHI is directly related to DNI by the cosine of theta. Thus, at a given level of GHI, a higher DNI could indicate a lower solar angle of
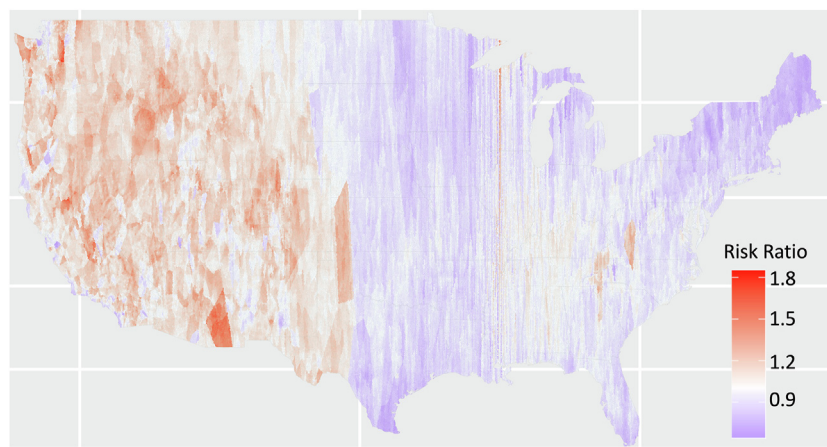
Figure 3. Plot of active exudative age-related macular degeneration (AMD) risk ratio across the contiguous United States. Active exudative AMD risk for 75-year-old White, pseudophakic woman with Medicare insurance calculated at each location of global horizontal irradiance (GHI) measurement, normalized by the mean risk taken over all GHI locations.

incidence (larger theta away from vertical). In this case, the light would pass through a greater volume of atmosphere. "Rayleigh scatter" of light in the atmosphere is caused by particles smaller than the wavelength of light and tends to scatter shorter wavelengths more than longer ones.[30] A lower GHI-to-DNI ratio might mean that lower frequencies are preferentially reaching the retina, which may protect against AMD risk. Our hypothesized mechanism for the effects of latitude, vitamin D production in skin (discussed next), contrasts with this explanation, and prior work has suggested a U-shaped, nonmonotonic association of light exposure with AMD risk.[13] Alternatively, at a given GHI, a higher DNI could imply a lower DHI and might mean more direct light is incident on skin. Perhaps direct light induces vitamin D synthesis more effectively than scatter irradiance.[31]

Other potential mechanisms rely on human interaction with the environment. Global horizontal irradiance may better indicate the irradiance incident on the human retina than DNI does. Human eyes are situated to naturally view the horizon, and typically the light incident on the retina is scatter irradiance, rather than light directly from the solar disc. Thus, higher GHI may increase AMD risk because it indicates the amount of incident light on the retina. In contrast, higher DNI might protect against AMD by altering behavior. Stronger direct sunlight may encourage sun-avoidance behaviors, such as wearing sunglasses or hats to reduce the strength of incoming rays and eliminate direct sunlight from entering the eye. Ultimately, further study is necessary to establish any mechanisms underlying these associations.

Latitude and AMD Risk

Our results suggest higher latitude is associated with an increased risk of developing active exudative AMD. Prior work has suggested similar findings. A 2016 meta-analysis by Reibaldi et al[32] demonstrated a protective association of lower latitude and increased solar insolation on both early and late AMD prevalence. They hypothesized

vitamin D levels may mediate the protective effects. Endogenous production of vitamin D in the skin as a response to UV light exposure is the primary source of vitamin D in humans.[33] A 2018 study by Kim and Park[34] demonstrated serum vitamin D deficiency is correlated with late-stage AMD risk, and they hypothesize that low vitamin D may specifically increase risk of exudative AMD. Incident sunlight rays at higher latitudes must travel through a greater distance of atmosphere than at the equator.[35] Additionally, atmosphere at higher latitudes has a higher concentration of UV-absorbing ozone than the atmosphere in the tropics.[35] Combined, these effects can cause higher latitude locations to receive less UV radiation, which could be a mechanistic effect in increasing AMD risk at higher latitudes. We observed correlations in our data among GHI, DNI, and latitude. However, this does not suggest that the observed associations of GHI and DNI with active exudative AMD are due to their correlation with latitude. Because we did not use any regularization in our models, we expect our multivariate regression analysis to identify the individual, "conditional," contributions of each of our variables. Generally, examining the mechanism of the effects of light on AMD risk is especially challenging because sunlight incident on skin may have different effects than sunlight incident on the retina or other ocular structures.

Environmental Pollutants and AMD Risk

Environmental pollutants increase the risk of developing cardiovascular disease, lung cancer, and other diseases.[36—38] We found that environmental pollution also contributes to the risk of developing AMD; 2 pollutants, ozone and nitrogen dioxide, significantly increased this risk. The deleterious effects of these pollutants arise, in part, from inducing oxidative stress in affected tissues. This mechanism may also contribute to disease progression in AMD, and the findings presented here highlight several promising directions for further research aimed at identifying the mechanisms that contribute to AMD.

US Risk-Prediction Map

The risk map shown in Figure 3 accounts for the influence of all environmental variables used in our model for active exudative AMD versus all other AMD. Risk ratio is plotted to illustrate differences with respect to the mean calculated risk. The actual risk values are based on an example patient (75-year-old White, pseudophakic female with Medicare insurance), and using different patient characteristics would shift both the mean risk and risk at each location of measurement across the United States. The overall distribution of risk can be more finely understood by examining both the distribution of variables across the United States and the isolated influence of subsets of variables on risk. To facilitate this, the values for environmental variables used in the model are plotted across the United States in Figure S2 (available at www.ophthalmologyscience.org), and the influence on risk of 9 different subsets of variables is shown in Figure S5. An east-west division is apparent, with generally higher risks west of the Rocky Mountains (with other pockets of high risk around Tennessee and the Appalachian Mountains). Looking at Figure S5, solar, elevation, and latitude variables appear to drive much of this east-west division of risk, with pollution also contributing. Comparing the environmental variable subsets "Temperature and Precipitation" and "All Except Temperature and Precipitation," it appears temperature and precipitation variables contribute to the higher risk seen along the southwestern US border. It should be noted that this model demonstrates association between risk and environmental variables but does not imply causality.

Several other artifacts in this risk map can also be explained. The large, shaded polygons arise from the low-resolution measurement of pollution variables (Fig S2). For example, the higher-risk, red quadrilateral in southern Arizona appears to arise from high levels of sulfur dioxide. The variables measured at higher resolution (e.g., GHI and DNI) yield smoother maps. Another artifact is north-south streaks observed running from the Upper Peninsula of Michigan to the Mississippi River delta. The risk maps are made at the resolution of GHI and DNI measurements, and these streaks arise due to the locations of these measurements. Looking at the "Temperature and Precipitation" variable subset map, the high-risk north-south streaks appear to be due to values of temperature and precipitation. These values may have measurement errors in the source data, but we are unable to definitively identify artifacts in the public data sources. These approaches to mapping risk across the United States are useful for generating hypotheses regarding disease mechanisms.

Strengths and Limitations

Our database retrospective cohort design allowed us to efficiently analyze a large cohort of patients across the United States and have sufficient power to analyze AMD-risk associations within AMD subgroups. The 9 884 527 patients included in our patient cohort is larger than in former studies,[14,32] and we obtained data at the individual patient level, rather than obtaining aggregated data at the level of studies, as can be the case for meta-analyses. This allowed us to control for phakic status and smoking history, in addition to demographic variables. The use of public environmental data and the open-sourcing of this organized data will enable future research to extend our approach, either by investigating associations of this data with other ocular diseases or by expanding the environmental factors taken into account.

Study Limitations

This study has several limitations. Because of the retrospective database cohort design, we used a coarse analysis of environmental exposures. Patients did not have lifetime location information or information about exposure mitigation (e.g., sunlight avoidance, occupational exposure, time spent outdoors) and were instead attributed the environmental factor values from the single location to which they were assigned (the location of the first visited provider from 2016 to 2018 with postal-code information). Our measures of solar irradiance do not distinguish the ocular exposure level from the general levels of irradiance at a geographic location. Because the IRIS Registry is built from ophthalmology practices, our analyses are performed with respect to a population of patients receiving ophthalmic care, rather than the general population of the United States. Additionally, because of limitations in the data, we used intravitreal injection CPT codes to help identify clinically active exudative AMD. This would fail to capture patients who have active exudative disease and refuse treatment, although we expect these patients to be a minority. Because of the retrospective and associational nature of our study, causal effects of environmental influences on AMD development cannot be established.

Conclusions

This study advances the understanding of environmental influences on AMD risk at different stages of the disease, highlighting associations among latitude, solar irradiance, and pollution and active exudative AMD. By merging public environmental data sources with the IRIS Registry, we achieve an analysis at a scale that would be prohibitive using prospective, traditional methods of data collection. The influence of natural environmental factors on AMD risk has been debated in the field, and our inclusion of approximately 10 million patients allowed us to investigate these questions across the United States in a large sample and within disease subgroups. Our findings indicate different stages and complications of AMD may differ in their etiologies. Although we hypothesize about the mechanisms driving the observed associations, future studies will be needed to investigate the precise causes involved. These studies might include additional database studies and in vitro experimental studies. Extensions of this work could involve analysis of additional interesting environmental and geographic factors and their influence on AMD or other ocular diseases.

Footnotes and Disclosures

References

1. Resnikoff S, Pascolini D, Etya'ale D, et al. Global data on visual impairment in the year 2002. *Bull World Health Organ*. 2004;82:844−851.
2. Lim LS, Mitchell P, Seddon JM, et al. Age-related macular degeneration. *Lancet*. 2012;379:1728−1738.
3. Heiba IM, Elston RC, Klein BE, Klein R. Sibling correlations and segregation analysis of age-related maculopathy: the Beaver Dam Eye Study. *Genet Epidemiol*. 1994;11:51−67.
4. Meyers SM. A twin study on age-related macular degeneration. *Trans Am Ophthalmol Soc*. 1994;92:775−843.
5. Seddon JM, Ajani UA, Mitchell BD. Familial aggregation of age-related maculopathy. *Am J Ophthalmol*. 1997;123:199−206.
6. The Eye Disease Case-Control Study Group. Risk factors for neovascular age-related macular degeneration. *Arch Ophthalmol*. 1992;110:1701−1708.
7. Age-Related Eye Disease Study Research Group. Risk factors associated with age-related macular degeneration. A case-control study in the age-related eye disease study: Age-Related Eye Disease Study Report Number 3. *Ophthalmology*. 2000;107:2224−2232.
8. Klein R, Clegg L, Cooper LS, et al. Prevalence of age-related maculopathy in the Atherosclerosis Risk in Communities Study. *Arch Ophthalmol*. 1999;117:1203−1210.
9. Klein R, Klein BEK, Tomany SC, Moss SE. Ten-year incidence of age-related maculopathy and smoking and drinking: the Beaver Dam Eye Study. *Am J Epidemiol*. 2002;156:589−598.
10. Armstrong RA, Mousavi M. Overview of risk factors for age-related macular degeneration (AMD). *J Stem Cells*. 2015;10:171−191.
11. Chakravarthy U, Wong TY, Fletcher A, et al. Clinical risk factors for age-related macular degeneration: a systematic review and meta-analysis. *BMC Ophthalmol*. 2010;10:31.
12. Rudnicka AR, Kapetanakis VV, Jarrar Z, et al. Incidence of late-stage age-related macular degeneration in American Whites: systematic review and meta-analysis. *Am J Ophthalmol*. 2015;160:85−93.e3.
13. Delcourt C, Cougnard-Grégoire A, Boniol M, et al. Lifetime exposure to ambient ultraviolet radiation and the risk for cataract extraction and age-related macular degeneration: the Alienor Study. *Invest Ophthalmol Vis Sci*. 2014;55:7619−7627.
14. Zhou H, Zhang H, Yu A, Xie J. Association between sunlight exposure and risk of age-related macular degeneration: a meta-analysis. *BMC Ophthalmol*. 2018;18:331.
15. Chiang MF, Sommer A, Rich WL, et al. The 2016 American Academy of Ophthalmology IRIS® Registry (Intelligent Research in Sight) Database: characteristics and methods. *Ophthalmology*. 2018;125:1143−1148.
16. United States Census Bureau. ZIP Code Tabulation Areas (ZCTAs). Available at: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html#:~:text=ZIP%20Code%20Tabulation%20Areas%20(ZCTAs)%20are%20gener-alized%20areal%20representations%20of,station%20associated%20with%20mailing%20addresses. Accessed May 1, 2021.

17. United States Census Bureau. Glossary. Available at: https://www.census.gov/programs-surveys/geography/about/glossary.htmlp. Accessed May 1, 2021.
18. Kyle Walker BR. *Package "tigris."*. R; 2020. https://cran.r-project.org/web/packages/tigris/tigris.pdf, 1 May 2021.
19. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. 2011. *Journal of Machine Learning Research*. 2011;12:2825−2830.
20. United States Geological Survey. LidarExplorer. Available at: https://prd-tnm.s3.amazonaws.com/LidarExplorer/index.html#. Accessed April 16, 2021.
21. United States Geological Survey. About 3DEP Products & Services. USGS. Available at: https://www.usgs.gov/core-science-systems/ngp/3dep/about-3dep-products-services?qt-science_-support_page_related_con=0#qt-science_support_page_related_con. Accessed May 1, 2021.
22. United States Department of Energy. NSRDB Data Viewer. NSRDB: National Solar Radiation Database. Available at: https://maps.nrel.gov/nsrdb-viewer/?aL=x8CI3i%255Bv%255D%3Dt%26VRLt_G%255Bv%255D%3Dt%26VRLt_G%255Bd%255D%3D1%26ozt_aP%255Bv%255D%3Dt%26ozt_aP%255Bd%255D%3D2&bL=clight&cE=0&lR=0&mC=4.740675384778373%2C23.203125&zL=2. Accessed April 15, 2021.
23. Vaisala Energy. What is Global Horizontal Irradiance? Vaisala Energy Support. Available at: https://www.3tier.com/en/support/solar-prospecting-tools/what-global-horizontal-irradiance-solar-prospecting/. Accessed May 6, 2021.
24. United States Department of Energy. U.S. Data. NSRDB: National Solar Radiation Database. Available at: https://nsrdb.nrel.gov/about/u-s-data.html#psm. Accessed May 1, 2021.
25. National Oceanic and Atmospheric Administration. 1981-2010 U.S. Climate Normals. Available at: https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/climate-normals/1981-2010-normals-data. Accessed April 16, 2021.
26. United States Environmental Protection Agency. Pre-Generated Data Files. Available at: https://aqs.epa.gov/aqsweb/airdata/download_files.html#Annual. Accessed November 29, 2021.
27. United States Environmental Protection Agency. NAAQS Table. Available at: https://www.epa.gov/criteria-air-pollutants/naaqs-table. Accessed February 16, 2022.
28. Wood SN, Goude Y, Shaw S. Generalized additive models for large datasets. *Journal of the Royal Statistical Society*, Series C 2015;64(1):139-155. https://rss.onlinelibrary.wiley.com/doi/full/10.1111/rssc.12068
29. Zipp K. *How Do You Measure Solar Energy? Solar Power World*; 2013. https://www.solarpowerworldonline.com/2013/04/how-do-you-measure-solar-energy/#:∼:text=Global%20Horizontal%20Irradiance%20(GHI)%2C,in%20units%20of%20W%2Fm2. Accessed May 4, 2021.
30. Humboldt State University. Atmospheric Scattering. GSP 216: Introduction to Remote Sensing. http://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson2-1/scatter.html. Accessed May 4, 2021.
31. Turnbull DJ, Parisi AV, Kimlin MG. Vitamin D effective ultraviolet wavelengths due to scattering in shade. *J Steroid Biochem Mol Biol*. 2005;96:431−436.
32. Reibaldi M, Longo A, Pulvirenti A, et al. Geo-epidemiology of age-related macular degeneration: new clues into the pathogenesis. *Am J Ophthalmol*. 2016;161:78−93.e1−2.
33. Nair R, Maseeh A. Vitamin D: the "sunshine" vitamin. *J Pharmacol Pharmacother*. 2012;3:118−126.
34. Kim KL, Park SP. Association between serum vitamin D deficiency and age-related macular degeneration in Koreans: clinical case-control pilot study. *Medicine*. 2018;97:e11908.
35. United States Environmental Protection Agency. *UV Radiation*. Environmental Protection Agency; 2010. https://www.epa.gov/sites/production/files/documents/uvradiation.pdf, 1 May 2021.
36. Bhatnagar A. Cardiovascular effects of particulate air pollution. *Annu Rev Med*. 2022;73:393−406.
37. Joshi SS, Miller MR, Newby DE. Air pollution and cardiovascular disease: the Paul Wood Lecture. *British Cardiovascular Society 2021*. *Heart*. 2022 Jan 24. https://doi.org/10.1136/heartjnl-2021-319844. heartjnl-2021-319844, Online ahead of print.
38. Quezada-Maldonado EM, Sánchez-Pérez Y, Chirino YI, García-Cuellar CM. Airborne particulate matter induces oxidative damage, DNA adduct formation and alterations in DNA repair pathways. *Environ Pollut*. 2021;287:117313.