Cureus
Part of SPRINGER NATURE

Open Access Original Article

Age, Race, Gender, Ethnicity, Preferred Language, and Marital Status as Factors Associated With Glaucoma Classification: A Retrospective Cross-Sectional Study

Review began 11/02/2026
Review ended 11/10/2026
Published 11/16/2026

© **Copyright** 2026
Helou et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.119799

Nancy Helou [1], Jinan Karameh [2]

1. Optometry and Healthcare Data Analytics, Independent Researcher, Rancho Cucamonga, USA 2. Optometry, Independent Researcher, Rancho Cucamonga, USA

Corresponding author: Jinan Karameh, jinankarameh@live.com

Abstract

Background

Glaucoma refers to a group of related conditions that damage the optic nerve and represent a major cause of irreversible blindness worldwide. It often progresses without any noticeable symptoms until the later stages, and a late diagnosis can lead to permanent visual impairment as well as higher healthcare costs. Demographic factors may be associated with differences in glaucoma classification and detection. This study examined associations between age, race, gender, ethnicity, preferred language, and marital status, with a special focus on nonlinear patterns related to age and interactions among demographic factors.

Methodology

This retrospective, cross-sectional secondary analysis included 15,165 patient records from the publicly available FairFedMed-Oph dataset. Demographic variables included age, gender, race, ethnicity, preferred language, and marital status. We used descriptive statistics, chi-square tests, and multivariable logistic regression to examine associations with glaucoma classification. We used a cubic spline logistic regression model to test for nonlinear age associations and explored age-by-race and age-by-gender interactions. Model fit and discriminatory performance were assessed using the Akaike information criterion (AIC), likelihood-ratio tests, and five-fold stratified cross-validation.

Results

Of the 15,165 patients included, 7,431 (49.0%) were classified as glaucoma-positive. The mean age of the glaucoma-positive patients was significantly higher than that of glaucoma-negative patients (64.77±16.49 years vs 57.90±15.38 years, p<0.001). The proportion classified as glaucoma-positive was higher among Black patients (n=1,435; 64.3%) than among Asian (n=620; 48.5%) and White (n=5,376; 46.1%) patients. Age was significantly associated with glaucoma classification in the expanded multivariable model (adjusted odds ratio (aOR), 1.029 per year; 95% confidence interval (CI), 1.027-1.032). Black patients had higher adjusted odds than Asian patients (aOR, 1.766; 95% CI, 1.522-2.048). Gender, preferred language, and marital status were also significantly associated with glaucoma classification after adjustment, whereas ethnicity was not. Cubic spline modeling provided a significantly better fit than linear-age modeling (Akaike information criterion (AIC), 19,374.47 vs. 19,787.31; p<0.001). The age-by-race interaction was significant (p<0.001), whereas the age-by-gender interaction was not significant (p=0.161). Spline modeling slightly improved discriminatory performance in five-fold cross-validation (mean area under the curve (AUC), 0.6765 vs. 0.6619).

Conclusions

Age, race, preferred language, marital status, and gender were associated with glaucoma classification in this clinical dataset. Nonlinear age modeling improved fit, and the association between age and glaucoma classification varied by race. However, the modest cross-validated AUC indicates limited discriminatory performance using demographic variables alone. These findings support further investigation of demographic differences in glaucoma classification and suggest that incorporating clinical and ophthalmic measurements may be necessary to improve predictive performance.

Categories: Other, Public Health, Ophthalmology
Keywords: cubic splines, demographic factors, glaucoma, glaucoma classification, race and ethnicity

Introduction

Glaucoma represents a leading cause of irreversible blindness and affects approximately 80 million people worldwide [1]. It is also associated with disability [2]. The global burden of glaucoma is expected to increase

substantially [1].

Both open-angle glaucoma and angle-closure glaucoma can be primary diseases. Secondary glaucoma can result from trauma, certain medications such as corticosteroids, inflammation, tumors, or conditions such as pigment dispersion or pseudoexfoliation [3]. Glaucoma refers to a group of optic neuropathies characterized by progressive degeneration of retinal ganglion cells. These are central nervous system neurons with cell bodies in the inner retina and axons in the optic nerve. Degeneration of these cells contributes to optic disc cupping and visual loss [3].

With the global population aging, the number of people aged 40-80 years with glaucoma is expected to rise from 76 million in 2020 to 111.8 million by 2040 [1]. As the prevalence increases, demographic factors become increasingly relevant; in the United States, for example, Black people are disproportionately affected [4]. Nevertheless, current prevalence estimates for glaucoma and vision-impairing glaucoma at national, state, and county levels remain limited [5].

Low awareness significantly limits timely diagnosis and treatment. Surveys indicate that only 10% to 50% of people with glaucoma are aware of their condition [6]. Because glaucoma can remain asymptomatic until advanced stages, the true number of affected individuals may be higher than reported [6]. More than half of glaucoma cases in the United States are undiagnosed, restricting opportunities for treatment and monitoring [5,7].

The mechanisms underlying the association between age and glaucoma risk remain insufficiently understood. The increased prevalence of glaucoma in older adults cannot be attributed solely to age-related elevations in intraocular pressure (IOP). This suggests that aging may increase the optic nerve's vulnerability to IOP-related damage and contribute to retinal ganglion cell (RGC) loss [8]. Given the increase in glaucoma prevalence with age, this association highlights the importance of investigating approaches to early detection and timely intervention among older adults to potentially reduce vision loss and disability [9].

Clinically, glaucoma includes several forms with different mechanisms and risks of visual loss. Based on anterior chamber angle configuration, glaucoma can be widely classified as open-angle or angle-closure. In the United States, more than 80% of cases are open-angle glaucoma; however, angle-closure glaucoma is responsible for a disproportionate number of cases of severe vision loss [10].

In the United States, substantial racial disparities exist in the prevalence of glaucoma, with higher rates observed among Black/African American and Hispanic populations compared with White/Caucasian individuals. Despite this, the rate of glaucoma diagnosis among Black/African Americans is lower than expected based on population-based prevalence estimates, and they are less likely than White/Caucasians to undergo visual field testing but more likely to require glaucoma surgery. This may indicate delayed access to care and/or greater clinical severity, calling for more urgent intervention [11].

Research suggests that men are more likely to develop a type of glaucoma called primary open-angle glaucoma. However, women have a higher risk of developing angle-closure glaucoma. It is not entirely clear why, and results may vary by study group. One thing to consider is that women tend to live longer than men, which means they have more time to develop glaucoma and potentially lose vision from it. This could be an important factor when considering the overall risk of glaucoma in men and women [12].

The social and demographic barriers that limit access to high-quality eye care, as well as their association with racial discrimination in glaucoma care, have not been thoroughly examined. Limited English proficiency may create barriers to accessing healthcare services and understanding health information. Studies have linked low health literacy and poor follow-up among patients with glaucoma to limited English proficiency [13].

Glaucoma costs the American economy $2.9 billion [14]. Recognizing and treating open-angle glaucoma (OAG) early is very important because disease progression could cause more serious health and cost problems [15]. As the disease progresses, vision loss increases and quality of life decreases [16].

Previous studies have examined individual demographic and socioeconomic variables linked to glaucoma. However, fewer studies have examined age, race, ethnicity, gender, preferred language, and marital status simultaneously in relation to glaucoma classification. Further research is needed to understand these demographic associations and determine whether they may inform future approaches to glaucoma detection and eye care.

This retrospective, cross-sectional study primarily explored associations between age, race, gender, ethnicity, preferred language, marital status, and glaucoma classification in the FairFedMed-Oph clinical dataset, after adjusting for these factors. These findings may inform future research on glaucoma identification and access to eye care. The secondary objectives were to assess nonlinear age associations, explore demographic interactions, and test the discriminatory performance of demographic models through

cross-validation. The outcome was binary glaucoma classification (glaucoma-positive versus glaucoma-negative), rather than glaucoma subtype, severity, or incidence. While this study does not directly assess screening effectiveness or health care costs, the results may provide a basis for further studies to identify ways to improve early detection, minimize preventable vision loss, and, perhaps, reduce the economic burden of glaucoma.

Materials And Methods

Study design and population

This retrospective cross-sectional secondary analysis used the publicly available FairFedMed-Oph dataset. The original dataset was created from de-identified clinical imaging data obtained from Massachusetts Eye and Ear (MEE), which is part of Harvard Medical School, between 2010 and 2022 [17].

The original investigators included one sample per patient, selected imaging from the most recent clinical visit, and randomly chose either the right or left eye. Eligible subjects had fundus and optical coherence tomography (OCT) scans with a signal-strength score of at least six. Board-certified ophthalmologists completed quality assessments of the images to determine their diagnostic value [17].

There were 16,681 patient samples in the original FairFedMed-Oph dataset. For this analysis, we used the publicly available glaucoma-detection subset, which included 15,165 patients. The analytical dataset was obtained from the publicly available FairFedMed-Oph glaucoma-detection data (data_summary.csv). We identified no missing values in the variables included in this analysis; however, some demographic information was recorded under the category 'Unknown'. The original investigators reported Institutional Review Board (IRB) approval and compliance with the Declaration of Helsinki [17]. The present study was a secondary analysis of publicly available, de-identified data and did not involve direct patient contact or access to identifiable patient information.

The FairFedMed-Oph dataset is available through the Harvard AI and Robotics repository on Hugging Face (https://huggingface.co/datasets/harvardairobotics/FairFedMed).

Definition of glaucoma

Glaucoma classification followed the method adopted by FairFedMed-Oph investigators. The classification was based on a reliable visual field examination performed within 30 days of fundus imaging [17]. To establish visual field reliability, fixation losses of 33% or less, false-positive rates of 20% or less, and false-negative rates of 20% or less were required. The original investigators defined glaucoma as a visual field mean deviation below -3 dB, together with abnormal results on both the glaucoma hemifield test and pattern standard deviation. Although fundus and optical coherence tomography (OCT) images were included in the original dataset, the published glaucoma classification criteria were based on visual field findings and did not specify additional structural or OCT thresholds. The original publication did not report a numerical threshold for abnormal pattern standard deviation. We did not modify the original binary glaucoma classification and used it as the outcome variable [17].

Study variables

The primary outcome was a binary glaucoma classification (glaucoma-positive: yes; glaucoma-negative: no). The primary explanatory variables were age, gender, and race. For regression analysis, we treated age as a continuous variable (years) and, for descriptive visualization, as a categorical variable. The age groups were younger than 40, 40-49, 50-59, 60-69, 70-79, and 80 years or older.

We analyzed gender, race, ethnicity, preferred language, and marital status as categorical variables. The first regression model included age, gender, and race. We also added ethnicity, preferred language, and marital status to the expanded model. We kept the Unknown categories for ethnicity, preferred language, and marital status as separate categories in the regression analysis rather than excluding them. Although we found no missing values in the variables analyzed, some demographic information was recorded as Unknown. We used the gender variable as recorded in the original dataset; the available documentation did not specify whether this variable represented sex assigned at birth or self-identified gender.

Statistical analysis

We used Python (Python Software Foundation (PSF), Beaverton, OR) and Google Colab (Google, Mountain View, CA) for statistical analysis and Microsoft Power BI (Microsoft, Redmond, WA) for visualization. The Python environment checked for this revision included Python 3.13.18, pandas 2.2.3, NumPy 2.4.6, SciPy 1.16.3, statsmodels 0.15.0, scikit-learn 1.6.1, and patsy 1.0.3. We used descriptive statistics (means, standard deviations, frequencies, and percentages) to summarize patient characteristics, and we reported the proportion of glaucoma-positive patients in each group. We compared glaucoma-positive and glaucoma-negative groups using Welch's independent-samples t-test for the continuous variable, age, and chi-square tests for the categorical variables. Effect sizes for chi-square tests were assessed using Cramér's V.

We tested the associations between demographic factors, including age, gender, race, ethnicity, preferred language, and marital status, and glaucoma using multivariable logistic regression. We reported adjusted odds ratios (aORs) with 95% confidence intervals (CIs).

We used logistic regression with linear and cubic B-spline models to examine the association between age and glaucoma classification. We used four degrees of freedom for the cubic spline, with one interior knot at the median observed age (64.01 years) and boundary knots at the minimum (8.95 years) and maximum (99.81 years) observed ages. The spline basis excluded an intercept, while the logistic regression models included an intercept. We compared both models using the Akaike information criterion (AIC) and likelihood-ratio tests. We performed a sensitivity analysis to examine the effect of the age range on the results by limiting the analysis to patients aged 40 years or older, as is typical in glaucoma epidemiological studies.

We used Python's statsmodels package to fit the logistic regression models. We encoded the categorical variables in the regression models, using female, Asian, Hispanic, English, and divorced as the reference categories for gender, race, ethnicity, preferred language, and marital status, respectively. The basic model included age, gender, and race, while the expanded model additionally included ethnicity, preferred language, and marital status. The expanded spline model replaced the linear age term with a cubic B-spline basis with four degrees of freedom. We tested age-by-race and age-by-gender interactions using linear age terms and compared models with and without interactions using likelihood-ratio tests. We also evaluated the age-by-race interaction using cubic B-spline age terms in the basic model. All evaluated basic and expanded logistic regression models converged successfully, including the linear-age and spline-age specifications and the additional spline age-by-race interaction models.

For the exploratory language analysis, we grouped preferred language into English and non-English categories. We excluded observations with Unknown language from this analysis only. We then calculated the proportion of glaucoma-positive observations by language group among White and Hispanic patients. These exclusions did not apply to the main regression models.

We used five-fold stratified cross-validation with shuffling and a random seed of 42. We fitted each model on the training data and evaluated it on the test data of each fold. For the spline models, we estimated the spline basis using the training data in each fold. We capped test-set ages outside the training-set boundaries at the corresponding boundary values before prediction. For each fold, we computed the area under the receiver operating characteristic curve (AUC) and reported the average AUC across the five folds. We considered $p<0.05$ statistically significant.

Results

Study population and demographic characteristics

The analytical dataset included 15,165 observations: 7,431 (49.0%) were glaucoma-positive, and 7,734 (51.0%) were glaucoma-negative. We found no missing values for the variables analyzed, and no exact duplicate records or duplicate filenames when we validated the dataset. The proportion of glaucoma-positive observations reflects the composition of this clinical dataset and should not be interpreted as population prevalence.

The mean age of observations with glaucoma was significantly higher than that of observations without glaucoma (64.77±16.49 vs 57.90±15.38 years, respectively; $p < 0.001$). Female observations accounted for 57.0% of the overall sample (8,645/15,165), while male observations accounted for 43.0% (6,520/15,165). In the unadjusted analysis, gender was not statistically associated with glaucoma classification ($\chi^2(1)=2.342$, $p=0.126$, Cramér's $V=0.012$).

There was a significant difference among the racial groups in terms of glaucoma classification ($\chi^2(2)=248.308$, $p<0.001$, Cramér's $V=0.128$). Black observations in the clinical dataset had the highest proportion of glaucoma-positive cases (1,435/2,231; 64.3%), followed by Asian observations (620/1,278; 48.5%) and White observations (5,376/11,656; 46.1%). Furthermore, language was significantly associated with glaucoma classification ($\chi^2(3)=199.269$, $p<0.001$, Cramér's $V=0.115$), as was marital status ($\chi^2(5)=215.834$, $p<0.001$, Cramér's $V=0.119$), but ethnicity was not ($\chi^2(2)=5.332$, $p=0.0695$, Cramér's $V=0.019$). Table *1* provides detailed demographic group characteristics for each group.

Characteristic	Overall (N=15,165)	Glaucoma-negative (n=7,734)	Glaucoma-positive (n=7,431)	P-value
Age, mean±SD (years)	61.27±16.30	57.90±15.38	64.77±16.49	<0.001
Gender				0.126
Female	8,645 (57.0%)	4,456 (57.6%)	4,189 (56.4%)	
Male	6,520 (43.0%)	3,278 (42.4%)	3,242 (43.6%)	
Race				<0.001
Asian	1,278 (8.4%)	658 (8.5%)	620 (8.3%)	
Black	2,231 (14.7%)	796 (10.3%)	1,435 (19.3%)	
White	11,656 (76.9%)	6,280 (81.2%)	5,376 (72.3%)	
Ethnicity				0.0695
Hispanic	564 (3.7%)	261 (3.4%)	303 (4.1%)	
Non-Hispanic	14,065 (92.7%)	7,195 (93.0%)	6,870 (92.5%)	
Unknown	536 (3.5%)	278 (3.6%)	258 (3.5%)	
Language				<0.001
English	14,028 (92.5%)	7,381 (95.4%)	6,647 (89.4%)	
Other	607 (4.0%)	202 (2.6%)	405 (5.5%)	
Spanish	248 (1.6%)	76 (1.0%)	172 (2.3%)	
Unknown	282 (1.9%)	75 (1.0%)	207 (2.8%)	
Marital status				<0.001
Divorced	1,028 (6.8%)	474 (6.1%)	554 (7.5%)	
Legally separated	138 (0.9%)	57 (0.7%)	81 (1.1%)	
Married or partnered	8,741 (57.6%)	4,701 (60.8%)	4,040 (54.4%)	
Single	3,962 (26.1%)	2,041 (26.4%)	1,921 (25.9%)	
Unknown	365 (2.4%)	187 (2.4%)	178 (2.4%)	
Widowed	931 (6.1%)	274 (3.5%)	657 (8.8%)	

TABLE 1: Demographic Characteristics of the Study Population by Glaucoma Classification

Values are n (%) unless otherwise indicated. Percentages use the column total as the denominator. Age was compared using Welch's t-test; categorical variables were compared using Pearson's chi-square test. P-values apply to the overall characteristic, not individual categories. SD, standard deviation.

Age, race, and ethnicity distribution of glaucoma classification

The distribution of glaucoma-positive observations across age categories is presented in Figure *1*.

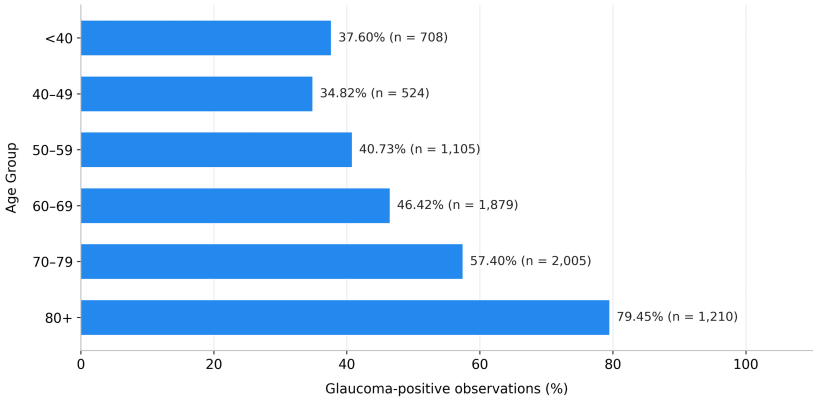


FIGURE 1: Proportion of Glaucoma-Positive Observations by Age Group

Percentages represent the number of glaucoma-positive observations divided by the total number of observations within each age category.

Figure *2* illustrates the corresponding distribution across racial and ethnic subgroups.

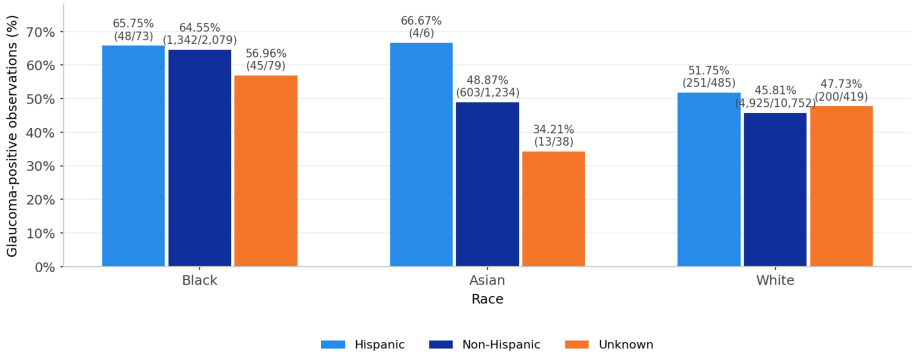


FIGURE 2: Proportion of Glaucoma-Positive Observations by Race and Ethnicity

Percentages were calculated by dividing the number of glaucoma-positive observations by the total number of observations within each race–ethnicity subgroup.

Multivariable logistic regression analysis

A model that included multiple variables (age, gender, and race) showed that age was associated with glaucoma classification (aOR per year, 1.029; 95% CI, 1.027-1.032; p<0.001). Compared with Asian observations, Black observations had significantly higher odds of glaucoma classification (aOR, 1.801; 95% CI, 1.560-2.079; p<0.001), while White observations had significantly lower odds (aOR, 0.767; 95% CI, 0.681-0.865; p<0.001). In this first model, gender was not significantly associated with glaucoma classification (aOR for male versus female, 1.045; 95% CI, 0.977-1.116; p=0.199).

We also created a model that included ethnicity, language, and marital status as additional variables. The Akaike information criterion (AIC) of the expanded model was lower than that of the original model (AIC=19,787.31 vs. 20,020.41), and the likelihood-ratio test indicated a statistically significant improvement in model fit (χ^2=253.10, degrees of freedom [df]=10, p<0.001).

In the expanded model, increasing age remained statistically associated with glaucoma classification (aOR per year, 1.029; 95% CI, 1.027-1.032; p<0.001). Modestly elevated odds were observed for male observations versus female observations (aOR, 1.118; 95% CI, 1.045-1.197; p=0.0013). Using Asian observations as the reference category, Black observations had higher adjusted odds (aOR, 1.766; 95% CI, 1.522-2.048; p<0.001), whereas White observations had lower adjusted odds (aOR, 0.823; 95% CI, 0.726-0.933; p=0.0023).

After adjustment for the other demographic factors mentioned, language and marital status were significantly associated with glaucoma classification. Spanish-language observations had higher adjusted odds of glaucoma classification (aOR, 2.192; 95% CI, 1.569-3.062) than English-language observations, and observations conducted in other languages also had higher adjusted odds than those conducted in English (aOR, 1.786; 95% CI, 1.490-2.142). Compared with divorced observations, widowed observations had higher adjusted odds (aOR, 1.496; 95% CI, 1.233-1.814), while married or partnered observations had lower adjusted odds (aOR, 0.784; 95% CI, 0.686-0.897). Table *2* presents the complete expanded-model estimates, including nonsignificant categories.

Variable	Adjusted OR (95% CI)	P-value
Age (per year)	1.029 (1.027-1.032)	<0.001
Gender		
Female	Reference	-
Male	1.118 (1.045-.197)	0.0013
Race		
Asian	Reference	-
Black	1.766 (1.522-2.048)	<0.001
White	0.823 (0.726-0.933)	0.0023
Ethnicity		
Hispanic	Reference	-
Non-Hispanic	0.853 (0.687-1.058)	0.1484
Unknown	0.859 (0.652-1.133)	0.2828
Language		
English	Reference	-
Other	1.786 (1.490-2.142)	<0.001
Spanish	2.192 (1.569-3.062)	<0.001
Unknown	2.430 (1.840-3.208)	<0.001
Marital status		
Divorced	Reference	-
Legally separated	1.088 (0.748-1.584)	0.659
Married or partnered	0.784 (0.686-0.897)	0.0004
Single	1.110 (0.960-1.284)	0.1589
Unknown	0.913 (0.712-1.171)	0.4733
Widowed	1.496 (1.233-1.814)	<0.001

TABLE 2: Multivariable Logistic Regression Analysis of Demographic Factors Associated With Glaucoma Classification

Note. OR=odds ratio; CI=confidence interval. The expanded linear-age model includes age, gender, race, ethnicity, language, and marital status simultaneously. Reference categories are indicated. N=15,165 observations. The model intercept is not shown. Percentages are not presented because the estimates are adjusted odds ratios.

To assess whether the demographic associations remained consistent after accounting for nonlinear age

effects, we examined the expanded logistic regression model incorporating cubic B-splines for age. Male gender remained significantly associated with higher odds of glaucoma classification (aOR=1.106; 95% CI: 1.032-1.184; p=0.004). Compared with Asian participants, Black participants had higher adjusted odds (aOR=1.883; 95% CI: 1.622-2.187; p<0.001), whereas White participants had lower adjusted odds (aOR=0.808; 95% CI: 0.712-0.916; p<0.001). Compared with English-speaking participants, Spanish-speaking participants had higher adjusted odds (aOR=2.412; 95% CI: 1.722-3.378; p<0.001), as did participants reporting other languages (aOR=1.753; 95% CI: 1.458-2.109; p<0.001). Married or partnered participants had lower adjusted odds than divorced participants (aOR=0.782; 95% CI: 0.682-0.896; p<0.001). Ethnicity was not significantly associated with glaucoma classification in this model. Unlike in the expanded linear-age model, widowed status was not significantly associated with glaucoma classification after nonlinear age adjustment (aOR=1.093; 95% CI: 0.894-1.337; p=0.386). The expanded spline-age model converged successfully and had an AIC of 19,374.47.

Nonlinear association between age and glaucoma classification

We compared cubic spline logistic regression models with models that included age as a linear predictor to explore a nonlinear relationship between age and glaucoma classification.

In the initial model, the spline specification had a lower AIC (19,516.14) than the linear-age specification (20,020.41), and the likelihood-ratio test indicated a statistically significant improvement in model fit (χ^2=510.27; df=3; p<0.001).

The expanded model, which controlled for gender, race, ethnicity, language, and marital status, retained this improvement. The cubic spline-age model had a lower AIC than the corresponding linear-age model (19,374.47 versus 19,787.31) and showed a statistically significant improvement in fit on the likelihood-ratio test (χ^2=418.84; df=3; p<0.001). These results indicate a nonlinear relationship between age and glaucoma classification in the sample included in this study.

Additional demographic associations

There was a significant association between marital status and classification of glaucoma (χ^2=215.83; df=5; p<0.001; Cramér's V=0.119). Among participants with known marital status, married or partnered participants had the lowest glaucoma-positive proportion (n=4,040; 46.22%), followed by single participants (n=1,921; 48.49%), divorced participants (n=554; 53.89%), legally separated participants (n=81; 58.70%), and widowed participants (n=657; 70.57%).

When compared by racial and ethnic groups, descriptive analyses showed that there were differences in the percentage of glaucoma-positive individuals between the English-speaking and non-English-speaking groups. The percentage of individuals with glaucoma was higher among White non-English speakers (n=292/470; 62.13%) than among White English speakers (n=5,012/11,070; 45.28%). The percentages were 43.36% (n=147/339) among Hispanic English speakers and 69.37% (n=154/222) among Hispanic non-English speakers. We included Spanish and other reported languages in non-English languages, but excluded unknown languages. Exploratory interaction analyses found no statistically significant interactions between language and race (χ^2=1.289; df=2; p=0.525) or language and ethnicity (χ^2=0.812; df=1; p=0.367).

Interaction and sensitivity analyses

An age-by-race interaction term was statistically significant in the model adjusted for gender (likelihood-ratio χ^2=22.06; df=2; p<0.001), suggesting that the relationship between age and glaucoma classification differed by race. The age-by-gender interaction was not statistically significant in the model adjusted for race (likelihood-ratio χ^2=1.97; df=1; p=0.161).

The age-by-race interaction remained statistically significant after accounting for nonlinear age using cubic B-splines (likelihood-ratio $\chi^2(8)$=62.396, p<0.001). The model including the interaction had a lower AIC than the model without the interaction (19,469.74 vs. 19,516.14). Both models converged successfully.

Figure *3* shows the proportions of glaucoma-positive participants stratified by race, without adjustment. Black participants had the highest percentages in all age categories shown. In the 80 years and older age group, percentages were 89.5% (n=145/162) among Black participants, 87.2% (n=82/94) among Asian participants, and 77.6% (n=983/1,267) among White participants.

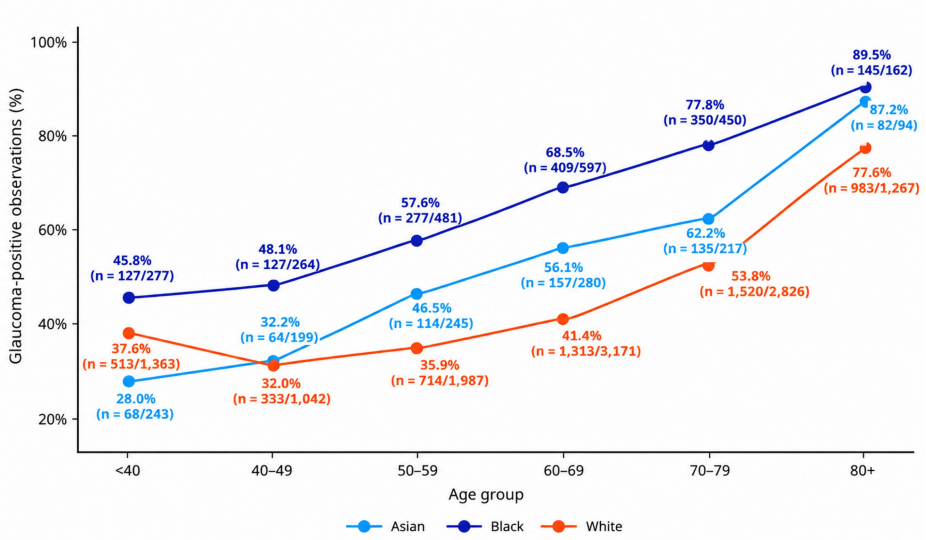


FIGURE 3: Glaucoma-Positive Proportion by Age Group and Race

The figure illustrates the percentage of glaucoma-positive observations across age groups stratified by race. Percentages are unadjusted and represent the proportion within each age-by-race subgroup.

Sensitivity analysis

A sensitivity analysis excluding younger observations (younger than 40 years) included 13,282 observations. In this subgroup, the cubic spline model adjusted for gender and race fit the data better than the corresponding linear-age model (AIC, 17,062.50 versus 17,212.13; likelihood-ratio χ^2=155.63; df=3; p<0.001). This suggests that the nonlinear age specification fits the data better for older observations.

Model validation and discriminatory performance

We compared the discriminatory performance of the expanded linear-age and cubic spline-age logistic regression models using cross-validation. Both models included age, gender, race, ethnicity, preferred language, and marital status. The only difference between the models was the functional form of age, which was modeled linearly in one model and using cubic splines in the other.

The linear-age model had a mean AUC of 0.6619, and the cubic spline-age model had a mean AUC of 0.6765, an absolute improvement of 0.0146.

The cubic spline model showed a better fit and modestly improved cross-validated discrimination, but its discriminatory performance remained generally limited. Table *3* reports the model fit statistics, interaction analyses, sensitivity results, and cross-validation results.

Analysis	Linear AIC / mean AUC	Spline AIC	LR χ^2 (df)	P-value
Model fit comparisons				
Initial model: linear vs spline age	20,020.41	19,516.14	510.27 (3)	<0.001
Expanded model: linear vs spline age	19,787.31	19,374.47	418.84 (3)	<0.001
Interaction analyses				
Age × race interaction	-	-	22.06 (2)	<0.001
Age × gender interaction	-	-	1.97 (1)	0.161
Sensitivity analysis (Age ≥40 years; n = 13,282)				
Age ≥40: linear vs spline age	17,212.13	17,062.50	155.63 (3)	<0.001
Five-fold stratified cross-validation				
Expanded linear-age model	0.6619	-	-	-
Expanded spline-age model	0.6765	-	-	-
Absolute AUC improvement (spline − linear)	0.0146	-	-	-

TABLE 3: Model Comparisons, Interaction and Sensitivity Analyses, and Cross-Validated Discriminatory Performance

Abbreviations: AIC, Akaike information criterion; AUC, area under the receiver operating characteristic curve; LR, likelihood-ratio; df, degrees of freedom.

Initial models included age, gender, and race. Expanded models additionally included ethnicity, language, and marital status. Spline models used a cubic spline specification for age.

Interaction analyses used models adjusted for the remaining gender or race covariate, as appropriate. The age ≥40 sensitivity analysis included age, gender, and race. Cross-validation compared expanded models; mean AUC values are shown in the second numeric column (not AIC). P-values refer to likelihood-ratio tests; cross-validated AUC values are descriptive and no significance test for AUC differences is reported.

Discussion

Principal findings

We analyzed demographic factors associated with glaucoma classification in 15,165 observations from the FairFedMed-Oph dataset. Age, race, preferred language, and marital status were significantly associated with glaucoma classification, and gender was significant following multivariate adjustment. Ethnicity was not significantly associated with glaucoma classification. There were racial differences within the analyzed clinical sample, with adjusted odds of glaucoma classification being higher in Black participants as opposed to Asian participants (aOR=1.766; 95% CI: 1.522-2.048).

One interesting finding was that the association between age and glaucoma classification was not linear. Cubic spline models fit significantly better than linear-age models ($p<0.001$), indicating that the relationship between age and glaucoma classification changes with age. The significant interaction between age and race ($p<0.001$) also suggested that the age effects varied by race. The descriptive patterns in Figure *3* further support these findings and reveal differences in glaucoma-positive classification proportions by age group and race.

Other demographic factors associated with glaucoma classification were marital status and preferred language. Married or partnered participants had lower adjusted odds of glaucoma classification than divorced participants. Preferred language was also significantly associated with glaucoma classification, with Spanish-speaking and other non-English-speaking participants having higher adjusted odds than English-speaking participants. However, the language-by-race and language-by-ethnicity interactions were not statistically significant, providing no significant evidence that the association between preferred language and glaucoma classification differed across these demographic groups. The results suggest further exploration of the potential influence of social support, health care access, and language barriers.

Although the models fit better, the improvement in cross-validated AUC was modest (0.6619 to 0.6765),

suggesting that demographic variables alone were not very discriminatory. Overall, these results highlight the need to include nonlinear age relationships and recognize that demographic variables alone have limited discriminatory ability for glaucoma classification.

Comparison with previous studies

Our results are consistent with previous studies that identified age as an important factor associated with glaucoma. In our study, glaucoma-positive observations were older than glaucoma-negative observations (mean age 64.77 years; glaucoma-negative mean age 57.90 years; p<0.001), and after multivariable adjustment, age remained significantly associated with glaucoma classification (aOR=1.029/year; 95% CI: 1.027-1.032). Ehrlich et al. reported that the prevalence of glaucoma rose from 1.09% of U.S. adults at age 40-64 years to 4.44% at age 65-79 and 7.75% at age 80 years and older [5].

Although our study focused on glaucoma classification, the analyses showed that glaucoma was more strongly associated with older ages, even though we did not assess prevalence. Our results also showed that the cubic spline model fit significantly better than the linear-age specification (p < 0.001), highlighting the relevance of considering nonlinear age relationships.

Our study also showed racial differences with the same directionality as previous epidemiological studies. The proportion of glaucoma-positive participants was higher among Black participants (n=1,435; 64.3%) than among Asian (n=620; 48.5%) and White (n=5,376; 46.1%) participants, and adjusted odds of glaucoma were higher among Black participants than among Asian participants (aOR=1.766; 95% CI: 1.522-2.048). Ehrlich et al. calculated the age- and sex/gender-standardized prevalence of glaucoma to be 3.15% (95% uncertainty interval (UI): 2.32%-4.09%) among non-Hispanic Black adults in the U.S., but only 1.42% (95% UI: 1.10%-1.85%) among non-Hispanic White adults [5]. In our study, after adjusting for other factors, ethnicity was not significantly associated with glaucoma classification. The variability in outcomes and study populations makes direct comparisons difficult, but overall, the results underscore the need to study racial/ethnic differences in glaucoma-related outcomes.

Acuff et al., in another study of 2,982 individuals with glaucoma from the All of Us Research Program, found that, on average, Black participants were diagnosed with glaucoma at a younger age than White participants (60 years vs. 66 years; p<0.001), and Hispanic participants were diagnosed at a younger age than non-Hispanic participants (61 years vs. 65 years; p=0.001) [18].

We did not assess age at diagnosis in our study, and ethnicity did not significantly differ for glaucoma classification after adjustment. Although these differing study populations, combined with different methods of racial and ethnic classification and outcome definitions, make direct comparisons difficult, these results underscore the importance of examining racial/ethnic differences in glaucoma-related outcomes.

Another significant finding was preferred language. Adjusting for these factors, participants who reported Spanish as their preferred language had higher odds of glaucoma classification compared to those who reported English as their preferred language (aOR=2.192; 95% CI: 1.569-3.062), and those who reported other languages had higher odds of glaucoma classification (aOR=1.786; 95% CI: 1.490-2.142). In a retrospective study of 29,891 patients, Halawa et al. found that 11.2% preferred a language other than English. They identified Black, Asian, and Hispanic patients as having higher glaucoma severity at presentation as compared to White or non-Hispanic patients. Adjusting for age, gender, and English proficiency, differences in severity of glaucoma at presentation were diminished, especially for Asian and Hispanic patients. The results suggest that language preference is associated with differences in glaucoma presentation and care [11]. In our study, we assessed glaucoma classification, not visual field grade. In addition, the exploratory language-by-race and language-by-ethnicity interactions were not significant (p=0.525 and p=0.367, respectively). Thus, although preferred language was a significant independent predictor of glaucoma classification in our analysis, we do not infer that language preference accounts for racial/ethnic differences in glaucoma classification.

For gender, male participants had slightly higher adjusted odds of glaucoma classification (aOR=1.118; 95% CI: 1.045-1.197), although this association was not significant in the unadjusted model (p=0.126). But the results of previous studies have been conflicting about sex-related differences in open-angle glaucoma. Some studies found a higher incidence of glaucoma among males, while others did not. Vajaranant et al. summarized evidence of increased glaucoma occurrence among men in some studies, and others showed no difference between men and women [19].

Finally, the results confirm that using only demographic variables to classify glaucoma has limited discriminatory ability. Although nonlinear age modeling improved model fit, the AUC increased only slightly, from 0.6619 to 0.6765, when cross-validated. Shaikh et al., in previous research, estimated that 78% of people with definite glaucoma in a population-based sample in the United States were previously undiagnosed and untreated, highlighting a problem in the detection of glaucoma [7].

While Shaikh et al. focused on undetected disease rather than prediction using demographic variables [7],

both studies showed the need for more comprehensive glaucoma assessments. Future models can be improved by adding clinical measurements, ophthalmic imaging, and healthcare-related variables.

Clinical and public health implications

Glaucoma is an important public health concern due to its potential for irreversible vision loss, impact on daily functioning, and economic costs to patients and healthcare systems. As vision worsens, patients may lose independence and quality of life and may have trouble reading, walking, climbing stairs, recognizing faces, and driving [20]. These limitations may also contribute to psychological problems, such as anxiety and depression. In a case-control study, patients with primary open-angle glaucoma (13.0% for anxiety and 10.9% for depression) had a higher prevalence of anxiety and depression than controls (7.0% and 5.2%, respectively; $p<0.05$) [21]. Glaucoma also significantly affects daily functioning. Over 40% of people with glaucoma reported difficulties with near-vision activities, and people with bilateral glaucoma were almost five times more likely to have severe near-vision difficulties than people without glaucoma. It can also make driving difficult. The Collaborative Initial Glaucoma Treatment Study found that over 50% of participants in the driving study had difficulty with glare and 22% had difficulty with peripheral vision tasks [21].

Beyond these clinical and functional impacts, glaucoma also carries high economic costs. Direct costs are associated with the patient, including medicines, diagnostic tests, clinical visits, and surgery, while indirect costs are associated with reduced productivity and income loss. The direct treatment cost estimates per patient per year in the United States range from $623 to $2,511, depending on disease severity. The disease's total cost to the nation is estimated at $12 billion by 2032 and $17.3 billion by 2050 [20].

The demographic associations found in this study emphasize the need for targeted glaucoma awareness, patient education, and timely ophthalmic referral due to the clinical and socio-economic implications. Targeting populations with well-known glaucoma susceptibility factors, such as older individuals, individuals with a family history of glaucoma, and underserved communities, could be a priority for awareness campaigns. Culturally sensitive, multilingual educational strategies and greater awareness of glaucoma among primary care and front-line health care professionals can complement these efforts by improving risk factor identification, patient communication, and prompt referral to a glaucoma specialist. These strategies can help overcome barriers to glaucoma care and disparities in access to health services.

Innovations such as artificial intelligence (AI) and teleglaucoma may improve early glaucoma detection and expand access to advanced eye care services. Optic disc photographs, OCT, and visual field analysis can be used with AI to detect suspicious glaucomatous changes, help clinical decision-making, and prioritize patients for specialist evaluation. AI in teleglaucoma services could help enable remote assessment, especially in underserved and rural areas, which may help to decrease geographic barriers, avoid unnecessary travel, and reduce delays in receiving ophthalmic care [14].

Teleglaucoma also has the potential to save on healthcare costs and improve accessibility. A Canadian cost-effectiveness study estimated that teleglaucoma screening would lower costs (CAD $872 per patient), reduce patient travel time by 61 hours, and reduce physician waiting time by 30% compared with in-person assessment. A 30-year simulation also projected a 24% reduction in glaucoma-related blindness, showing teleglaucoma's potential to improve accessibility and cost-effectiveness in glaucoma care [14].

Strengths, limitations, and further research

The main strengths of this study are that it included a large sample size, examined a variety of demographic factors, and applied nonlinear modeling and interaction analyses to examine associations related to glaucoma; however, there are factors to consider when interpreting results, including limited clinical and socioeconomic data, potential selection bias, and limited generalizability. Furthermore, because this study used a retrospective cross-sectional design, the observed associations cannot establish causal relationships between demographic factors and glaucoma classification.

The study used a selected, single-center clinical dataset with imaging results and reliable visual field examinations, and selected one eye per patient. Accordingly, the results may not reflect associations between classification and demographics for glaucoma in the general population. Other causes of visual field abnormalities may also influence the visual field classification. This analysis could not account for important clinical variables such as intraocular pressure, central corneal thickness, family history, or glaucoma subtype, which may have caused residual confounding.

No values were missing from the variables analyzed, but Unknown demographic categories represented unavailable demographic information. Preferred language does not directly measure English proficiency or communication barriers, and marital status does not directly measure social support. The expanded spline-age model confirmed significant associations for gender, race, preferred language, and married or partnered marital status, although some category-specific findings differed from the linear-age model. We evaluated internal model discrimination using five-fold cross-validation, but not calibration, external validity, or clinical usefulness.

The proportion with glaucoma was lowest among English-speaking participants in the racial/ethnic groups examined in the exploratory analysis. An important question that will need to be addressed in future research is how language-related communication barriers, access to health services, and socioeconomic factors help explain the racial/ethnic differences in glaucoma classification. Further research is warranted to investigate these relationships between marital status, preferred language, ethnicity, and glaucoma classification, particularly in Hispanic and non-Hispanic populations. It remains unclear whether these population differences reflect unequal underlying disease risk or unequal access to health care and diagnosis. Longitudinal studies should better explain these relationships by accounting for not only clinical risk factors but also socioeconomic factors.

Conclusions

Glaucoma is a major public health issue as it can be asymptomatic and may be diagnosed only at a late stage when substantial and irreversible damage has occurred. If left untreated, delayed diagnosis may result in visual impairments, trouble with daily living skills, loss of independence, and diminished quality of life. These complications remind us to detect glaucoma early, increase awareness of the disease, and maximize access to eye care.

This study found associations between age, race, preferred language, marital status, and gender and glaucoma classification within the selected clinical dataset, and it found a nonlinear association between age and glaucoma classification. However, these findings are exploratory and do not establish causality. Demographic variables alone showed limited ability to distinguish glaucoma-positive from glaucoma-negative observations. These results suggest that future research should evaluate whether combining demographic data with clinical data, ophthalmic imaging, and health care-related data improves predictive performance. Early detection and treatment can help prevent vision loss, disability, and the economic burden of glaucoma in the future.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Nancy Helou, Jinan Karameh

Acquisition, analysis, or interpretation of data: Nancy Helou, Jinan Karameh

Drafting of the manuscript: Nancy Helou, Jinan Karameh

Critical review of the manuscript for important intellectual content: Nancy Helou, Jinan Karameh

Supervision: Nancy Helou, Jinan Karameh

Disclosures

Human subjects: Informed consent for treatment and open access publication was obtained or waived by all participants in this study. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY: Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology. 2014, 121:2081-90. 10.1016/j.ophtha.2014.05.013
2. Ramulu P: Glaucoma and disability: which tasks are affected, and at what stage of disease? . Curr Opin Ophthalmol. 2009, 20:92-8. 10.1097/ICU.0b013e32832401a9
3. Weinreb RN, Aung T, Medeiros FA: The pathophysiology and treatment of glaucoma: a review . JAMA. 2014, 311:1901-11. 10.1001/jama.2014.3192
4. Muñoz B, West SK, Rubin GS, Schein OD, Quigley HA, Bressler SB, Bandeen-Roche K: Causes of blindness and visual impairment in a population of older Americans: the Salisbury Eye Evaluation Study. Arch Ophthalmol. 2000, 118:819-25. 10.1001/archopht.118.6.819
5. Ehrlich JR, Burke-Conte Z, Wittenborn JS, et al.: Prevalence of glaucoma among US adults in 2022. JAMA Ophthalmol. 2024, 142:1046-53. 10.1001/jamaophthalmol.2024.3884

6. Leite MT, Sakata LM, Medeiros FA: Managing glaucoma in developing countries . Arq Bras Oftalmol. 2011, 74:83-4. 10.1590/s0004-27492011000200001
7. Shaikh Y, Yu F, Coleman AL: Burden of undetected and untreated glaucoma in the United States . Am J Ophthalmol. 2014, 158:1121-1129.e1. 10.1016/j.ajo.2014.08.023
8. Jammal AA, Berchuck SI, Thompson AC, Costa VP, Medeiros FA: The effect of age on increasing susceptibility to retinal nerve fiber layer loss in glaucoma. Invest Ophthalmol Vis Sci. 2020, 61:8. 10.1167/iovs.61.13.8
9. Mitchell P, Smith W, Attebo K, Healey PR: Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. Ophthalmology. 1996, 103:1661-9. 10.1016/S0161-6420(96)30449-1
10. Friedman DS, Wolfs RC, O'Colmain BJ, et al.: Prevalence of open-angle glaucoma among adults in the United States. Arch Ophthalmol. 2004, 122:532-8. 10.1001/archopht.122.4.532
11. Halawa OA, Jin Q, Pasquale LR, et al.: Race and ethnicity differences in disease severity and visual field progression among glaucoma patients. Am J Ophthalmol. 2022, 242:69-76. 10.1016/j.ajo.2022.05.023
12. McMonnies CW: Glaucoma history and risk factors . J Optom. 2017, 10:71-8. 10.1016/j.optom.2016.02.003
13. Altangerel U, Nallamshetty HS, Uhler T, et al.: Knowledge about glaucoma and barriers to follow-up care in a community glaucoma screening program. Can J Ophthalmol. 2009, 44:66-9. 10.3129/i08-175
14. Thomas S, Hodge W, Malvankar-Mehta M: The cost-effectiveness analysis of teleglaucoma screening device . PLoS One. 2015, 10:e0137913. 10.1371/journal.pone.0137913
15. Prum BE Jr, Rosenberg LF, Gedde SJ, et al.: Primary Open-Angle Glaucoma Preferred Practice Pattern(®) Guidelines. Ophthalmology. 2016, 123:P41-111. 10.1016/j.ophtha.2015.10.053
16. Freeman EE, Muñoz B, West SK, Jampel HD, Friedman DS: Glaucoma and quality of life: the Salisbury Eye Evaluation. Ophthalmology. 2008, 115:233-8. 10.1016/j.ophtha.2007.04.050
17. Li M, Wen C, Tian Y, et al.: FairFedMed: benchmarking group fairness in federated medical imaging with FairLoRA. IEEE Trans Med Imaging. 2026, 45:1337-51. 10.1109/TMI.2025.3622522
18. Acuff K, Radha Saseendrakumar B, Wu JH, Weinreb RN, Baxter SL: Racial, ethnic, and socioeconomic disparities in glaucoma onset and severity in a diverse nationwide cohort in the United States. J Glaucoma. 2023, 32:792-9. 10.1097/IJG.0000000000002261
19. Vajaranant TS, Nayak S, Wilensky JT, Joslin CE: Gender and glaucoma: what we know and what we need to know. Curr Opin Ophthalmol. 2010, 21:91-9. 10.1097/ICU.0b013e3283360b7e
20. Allison K, Patel D, Alabi O: Epidemiology of glaucoma: the past, present, and predictions for the future . Cureus. 2020, 12:e11686. 10.7759/cureus.11686
21. Quaranta L, Riva I, Gerardi C, Oddone F, Floriani I, Konstas AG: Quality of life in glaucoma: a review of the literature. Adv Ther. 2016, 33:959-81. 10.1007/s12325-016-0333-6