BMJ Open Ophthalmology

Colombian Ocular Diseases Epidemiology Study (CODES): incidence and sociodemographic characterisation of keratoconus between 2015 and 2020

Germán Mejia-Salgado,[1,2] Carlos Cifuentes-González,[1] William Rojas-Carabali,[1] Laura Zarate-Pinzón,[1,2] Luisa Fernanda Peña-Pulgar,[1,2] Diego Polania,[1,2] Danna Lesley Cruz-Reyes,[3] Alejandra de-la-Torre [1,2]

ABSTRACT

Objective To estimate the incidence and describe the demographic characteristics of keratoconus (KC) in Colombia using national health registry data between January 1st 2015 and December 31th 2020.

Methods and analysis We conducted a nationwide, population-based study using the Integrated Social Protection Information System from the Colombian Ministry of Health, the unique official database in the country. We used the International Classification of Diseases code for KC (H186) to identify the number of new cases of KC and estimate the incidence rates both overall and according to age and sex. We made a standard morbidity ratio map to graph Colombia's morbidity risk of KC onset.

Results Of 50 372 424 subjects, 21 710 had KC between 2015 and 2020. However, due to the COVID-19 pandemic, all the incidence rates of this study were based on the 18 419 reported until 2019. The incidence rate in the general population was 10.36 (95% CI 10.08 to 10.64) per 100 000 inhabitants. The incidence peak among males was in their early 20s and females in their late 20s. The overall male to female incidence rate ratio was 1.60. Regarding the distribution of the disease, most cases were reported in Bogotá (48.64%), Antioquia (14.04%) and Cundinamarca (10.38%).

Conclusion We performed the first nationwide, population-based study of KC in Latin America, finding distribution patterns similar to those reported in the literature. This study provides valuable information on the epidemiology of KC in Colombia, which is helpful in the development of policies for the diagnosis, prevention and treatment of the disease.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The incidence of keratoconus has been described in other population-based studies, ranging from 1.3 to 32.3 per 100 000 inhabitants.
⇒ Data for South America are scarce and limited to cohorts that do not reflect nationwide epidemiology data.

WHAT THIS STUDY ADDS

⇒ This study describes the incidence of keratoconus in a developing country in South America based on a nationwide database for the first time, finding an incidence of keratoconus in Colombia of 10.36 (95% CI 10.08 to 10.64) per 100 000 inhabitants, with a high number of new cases among those in their 20s.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study describes and adds new information about the epidemiology of keratoconus in Latin American developing countries.
⇒ We provide valuable data for the development of policies for diagnosis, prevention and treatment of the disease.

INTRODUCTION

Keratoconus (KC) is a progressive ectatic corneal disease that leads to corneal protrusion and thinning, resulting in irregular astigmatism and visual impairment. Generally, it has been considered a bilateral but asymmetrical condition.[1–3] Clinically, it can be asymptomatic in subclinical stages until blindness in cases of scarring or hydrops.[1 3] Glasses and contact lenses can restore visual acuity in most patients with KC.[1 3] However, severe cases require surgical procedures such as cross-linking, intracorneal ring segments and corneal transplantation. Of all patients with KC, 10%–20% will undergo at least one penetrating keratoplasty.[4–6]

Although the aetiology of KC is not entirely understood, several risk factors have been associated with its development. Age, ethnicity, consanguinity, atopy and mechanical trauma are some of the main risk factors described.[7] However, environmental factors

such as ultraviolet exposure,[8–10] pollution[11 12] and smoking[13 14] have also been reported.

The number of KC cases has been increasing in the last decades. This could be explained by the improved understanding of the disease and new imaging modalities that can diagnose KC more frequently and at earlier stages than in the past.[15] The reported annual incidence ranges from 1.3 to 32.3 per 100 000 inhabitants.[16 17] Vast differences between studies may be attributed to study population, ethnicity, diagnostic criteria and diagnostic devices used.[7]

KC is more common in the Asian continent.[18] Thus, most studies related to the epidemiology and burden of the disease are focused on these populations, while few studies described the incidence and prevalence of KC in Latin America.[19 20] In Colombia, Barraquer-Coll *et al*[21] described a prevalence of KC and pellucid marginal degeneration of 2.84% in a single-centre, cross-sectional study including 91 426 patients. The Colombian Ocular Diseases Epidemiology Study (CODES) is a series of reports focusing on the description of epidemiology data for ocular diseases in Colombia. Herein, we report the incidence of KC using the national health registry data over a period of 6 years.

METHODS
Design
We conducted an observational, descriptive, claim database, cross-sectional study using the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.

Population
The information and data in this study were obtained from the national database created by the Colombian Ministry of Health, known as the System of Information of Social Protection (SISPRO), which stores, processes and systematises Colombian citizens' health registers (data set).[22] The data are collected by medical staff during each medical contact (inpatient or outpatient) from private and public health providers and insurers using the International Classification of Diseases (ICD-10). In addition, the demographic and clinical data are grouped in the RIPS (Individual Registry of Health Services Provision).[23]

It should be noted that the Colombian health system has one of the most prominent coverages in Latin America, encompassing 50 million inhabitants, representing 97.78% of the population in 2020, according to the last survey of the National Administrative Department of Statistics (DANE) (data set).[24]

Data collection
Data were extracted from the SISPRO dynamic tables from 2015 to 2020 (data set). The ICD-10 code for KC (H186) was used. To delimit the results, we performed the first filter by year (2015–2020) and the type of diagnosis (confirmed new) to determine the incidence. Afterwards, additional filters, such as residence location, age

and sex, were added to describe the sociodemographic status of the disease. Finally, a database in Microsoft Excel (Microsoft, Redmond, Washington, USA) was elaborated and validated to record the information.

Statistical analysis
The crude incidence rate of KC in the Colombian population was estimated for 6 years (2015–2020) using a standardised crude rate per 100 000 inhabitants of confirmed new KC cases divided by the consultation population of the Colombian medical service (considering these as the risk population). For the incidence analysis, the patients were divided and stratified by sex and age in quinquennium. Twenty-four patients were excluded from the analysis by sex because they did not have this information reported. Still, they were included in the analysis of the total patients by age. Considering that the data are secondary and may not represent the parameter of the total number of cases, 95% CI for the rates was calculated assuming a Poisson error. These analyses were performed with R V.4.0.4.

Additionally, since the pandemic significantly affected the results in 2020, we used a statistical model for spatial data based on the Gaussian Markov random fields (GMRF) to predict what would have happened this year regarding the incidence of the disease. We used the conditional autoregressive model, a type of GMRF. This model approach visualises the geographical domain as an undirected graph with a vertex in each region and an edge between two vertices if the corresponding regions share a geographical edge. This creates well-defined neighbours for each region, which are used to define the joint or conditional distribution. As a result, the distribution will be the multivariate normal distribution. More detailed information regarding the statistical analysis is found in online supplemental material 1.

Bias control
In this study, selection bias may occur due to several filters applied to the diagnostic data in SISPRO (data set). Also, due to the scarcity of cornea supraspecialists in the country, underestimation and misclassification bias with other ICD codes like H187 (other corneal deformities) may occur.[25] However, we only included patients with a 'confirmed new diagnosis' to prevent these biases. Additionally, the H186 ICD-10 code is not commonly used by general physicians but by ophthalmologists. Further, since we do not have access to clinical charts, we cannot assert that there are no repeated patients because it is secondary information. After applying filters, the total data accuracy cannot be guaranteed because the database does not allow us to rectify the identification numbers. However, the Colombian Ministry of Health and healthcare provider institutions constantly update and correct data with errors to control this bias. Previous studies based on SISPRO (data set) have determined the concordance rate of the database using ICD-10 with

medical records to be up to 83.4%.[26 27] Therefore, these data are the most accurate approximation.

RESULTS

The SISPRO database from 2015 to 2020 recorded 27 017 352–34 370 304 consultations (data set). From 2015 to 2020, there were 21 710 new cases of KC throughout the country. However, considering the high probability of under-reporting due to the COVID-19 pandemic in 2020, we analysed the 18 419 cases reported until 2019, including 9685 (52.58%) males and 8734 (47.42%) females.

There were 2600 (14.11%), 2604 (14.13%), 3293 (17.87%), 4165 (22.61%) and 5757 (31.25%) incident cases in 2015, 2016, 2017, 2018 and 2019, respectively. The incidence rate of KC in the general population was 10.36 (95% CI 10.08 to 10.64) cases per 100 000 inhabitants. In males, it was 12.76 (95% CI 12.37 to 13.14) cases per 100 000 inhabitants; in females, it was 7.96 (95% CI 7.60 to 8.32) cases per 100 000 inhabitants. The peak incidence was at 20–24 years of age in males (39.24; 95% CI 39.12 to 39.35) and at 25–29 years of age in females (19.32; 95% CI 19.19 to 19.44). Males had a higher incidence rate in all ages except for ages 5–9, 65–69 and

>80 years; the overall male to female ratio with regard to incidence was 1.60 (range: 0.66–2.35 in the different age groups). Detailed incidence results are presented in table 1 and figure 1.

The incidence of KC increased from 9.62 (95% CI 9.52 to 9.71) in 2015 to 14.02 (95% CI 13.95 to 14.08) in 2019. For 2020, we presented the real reported value of 8.65 (95% CI 8.56 to 9.73) and an estimated value based on the trend from the previous years (15.74; 95% CI 14.64 to 22.02), considering the impact of the COVID-19 pandemic on health assistance. The incidence per year can be found in figure 2.

Finally, regarding the distribution of the disease, most cases were reported in Bogotá (48.64%), Antioquia (14.04%) and Cundinamarca (10.38%). More detailed information is available in online supplemental material 2. It is important to note that in 9.61% of the patients, the region of origin was not reported. Figure 3 graphs the standard morbidity ratio maps to identify the morbidity of KC in each department per year; for this analysis, the 1771 cases with no department registered were excluded from the 18 419 cases of KC reported in 2019, for a total of 16 648 cases.

Table 1 Number of patients with newly diagnosed keratoconus and the estimated incidence rate (per 100 000 inhabitants) of keratoconus in Colombia in 2015–2019

Age group (years)	Total		Male		Female		Male to female ratio
	New cases (n)	Incidence* (95% CI)	New cases (n)	Incidence* (95% CI)	New cases (n)	Incidence* (95% CI)	
0–4	104	0.54 (0.06 to 1.02)	59	0.62 (0.09 to 1.14)	45	0.47 (0.02 to 0.91)	1.31
5–9	210	1.63 (1.29 to 1.96)	86	1.3 (1.07 to 1.52)	124	1.96 (1.51 to 2.40)	0.66
10–14	866	7.54 (7.37 to 7.71)	496	8.77 (8.64 to 8.89)	370	6.32 (6.06 to 6.57)	1.38
15–19	2386	20.57 (20.47 to 20.67)	1497	28.87 (28.75 to 28.98)	889	12.28 (12.11 to 12.44)	2.35
20–24	3269	28.59 (28.50 to 28.67)	1870	39.24 (39.12 to 39.35)	1399	17.94 (17.80 to 18.07)	2.18
25–29	3280	28.15 (28.06 to 28.24)	1784	36.99 (36.85 to 37.12)	1496	19.32 (19.19 to 19.44)	1.91
30–34	2589	23.84 (23.74 to 23.94)	1342	29.99 (29.82 to 30.15)	1247	17.7 (17.56 to 17.83)	1.69
35–39	1898	18.58 (18.46 to 18.69)	905	21.73 (21.51 to 21.94)	993	15.43 (15.27 to 15.58)	1.4
40–44	1172	13.05 (12.90 to 13.19)	537	14.85 (14.58 to 15.11)	635	11.25 (11.05 to 11.44)	1.32
45–49	838	9.23 (9.06 to 9.40)	354	9.79 (9.49 to 10.08)	484	8.68 (8.45 to 8.90)	1.12
50–54	645	6.83 (6.64 to 7.02)	272	7.12 (6.78 to 7.45)	373	6.55 (6.29 to 6.80)	1.08
55–59	500	5.85 (5.63 to 6.07)	212	6.07 (5.61 to 6.52)	288	5.64 (5.35 to 5.92)	1.07
60–64	280	3.84 (3.54 to 4.13)	117	3.85 (3.26 to 4.43)	163	3.83 (3.44 to 4.21)	1
65–69	183	3.06 (2.69 to 3.42)	70	2.81 (2.04 to 3.57)	113	3.31 (2.84 to 3.77)	0.84
70–74	94	2.06 (1.55 to 2.57)	41	2.07 (1.16 to 2.97)	53	2.06 (1.38 to 2.73)	1
75–79	62	1.77 (1.15 to 2.39)	29	1.97 (0.66 to 3.27)	33	1.58 (0.72 to 2.43)	1.24
>80	43	0.99 (0.24 to 1.73)	14	0.89 (0.84 to 0.93)	29	1.09 (0.18 to 1.99)	0.81
Overall	18 419	10.36 (10.08 to 10.64)	9685	12.76 (12.37 to 13.14)	8734	7.96 (7.60 to 8.32)	1.6

*Incidence=(number of confirmed new cases by quinquennial age during 2015–2019/registered SISPRO consultation by quinquennial age during 2015–2019)×100 000 inhabitants (data set).[22]
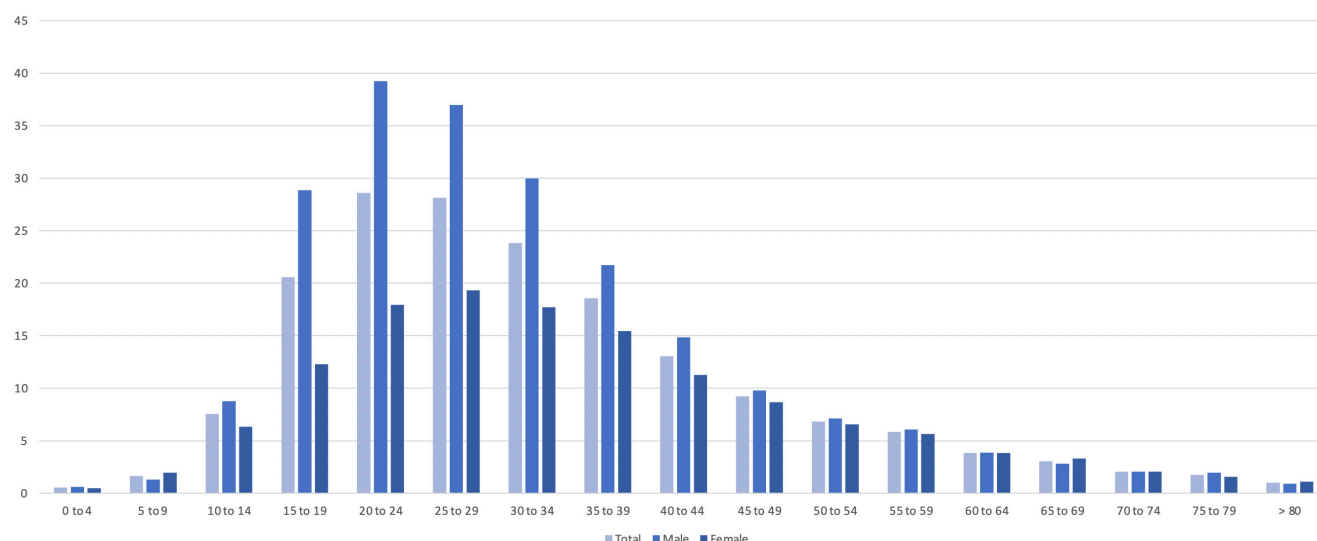SISPRO, System of Information of Social Protection.

Figure 1 Incidence rates of keratoconus both overall and according to age and sex. Incidence=(number of confirmed new cases by quinquennial age per year/SISPRO consultation registers by quinquennial age per year)×100 000 inhabitants (data set).[22] SISPRO, System of Information of Social Protection.

DISCUSSION

KC is an eye disease that, with proper treatment, can have excellent visual results; on the contrary, delays in its diagnosis and treatment can lead to blindness and requirement for surgical procedures.[4–6] Attention to KC has been increasing recently, and there has been a revolution in the knowledge required to diagnose KC. In the past, KC was diagnosed with slit lamp findings, refraction and keratometry.[28 29] However, corneal topography is now seen as the gold standard in the diagnosis and monitoring of ectatic disorders such as KC.[30] Studies focused on the epidemiology and incidence of KC offer valuable
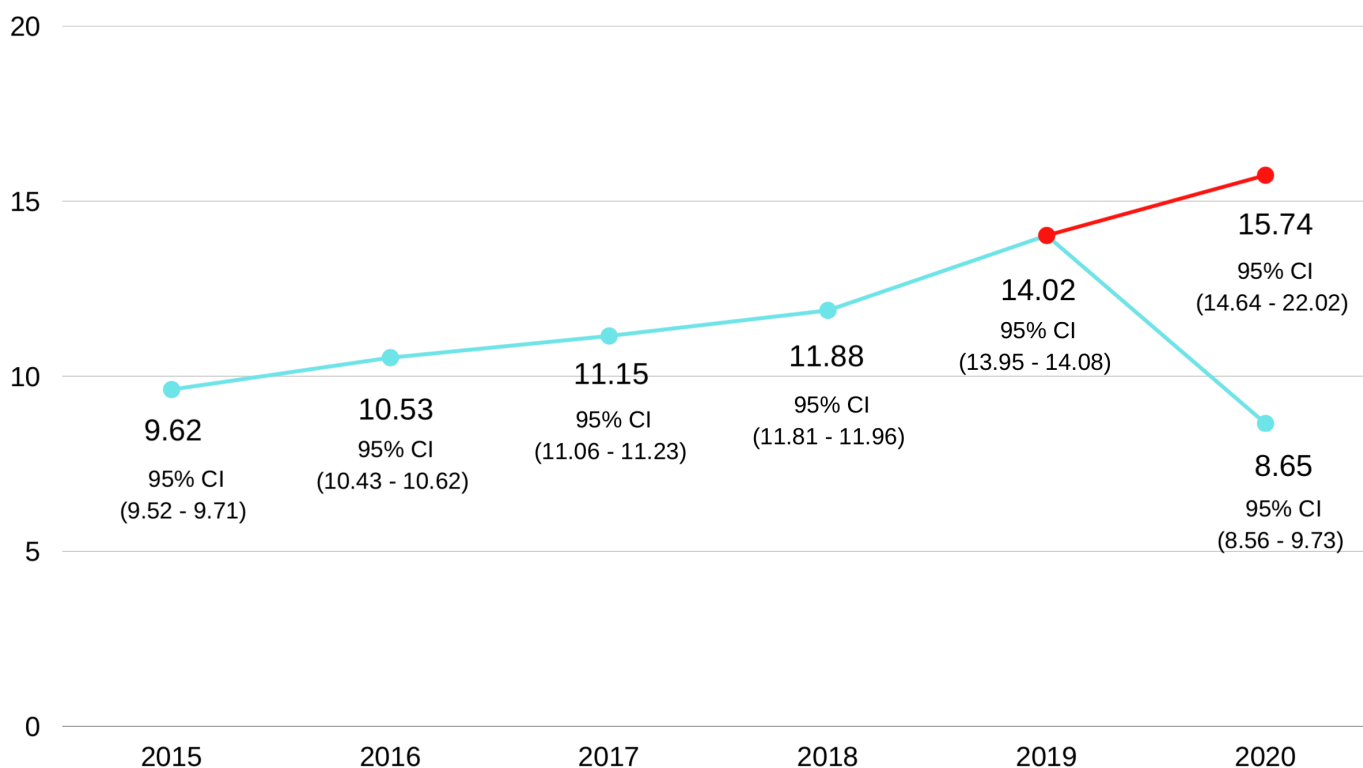

Figure 2 Annual distribution of keratoconus by quinquennial age in Colombia between 2015 and 2020. The blue line represents the incidence of keratoconus obtained by SISPRO data (data set). For 2020, the red line represents the estimated incidence based on SISPRO data analysis of the previous years, showing a difference of 7.09 per 100 000 inhabitants between the real and estimated incidence.[22] SISPRO, System of Information of Social Protection.
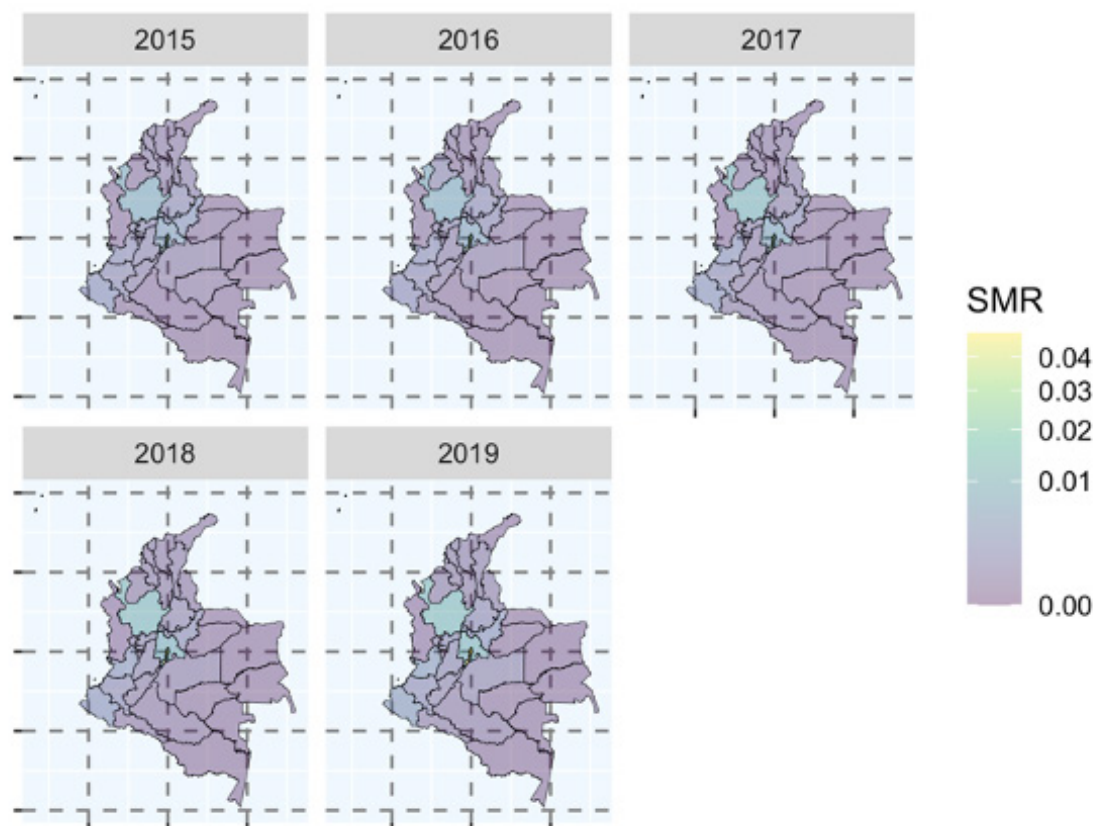
Figure 3 Standard morbidity ratio (SMR) maps of keratoconus in Colombia. The SMR of keratoconus was higher in the Andean region. Several factors could influence this higher risk; for example, the main cities are in these areas, and capital cities have an increased migration index from rural areas and more hospital access than other regions.

information for developing preventive and therapeutic strategies. We performed the first population-based study that describes the age-specific incidence of KC in Latin America.

Previously reported incidence rates of KC have varied widely depending on the study population, diagnostic criteria and geographical location. The incidence of KC in other population-based studies ranges from 1.3 to 32.3 per 100 000 inhabitants[16 17]; we report an incidence rate of KC in the general population of 10.36. This is lower than the incidence reported by Ziaei et al[31] of 22.3 per 100 000 inhabitants in a prospective, population-based survey during 2008–2009 in Iran, but higher than the incidence reported in Denmark and England of 3.6 and 3.5 per 100 000 person-years, respectively.[17 32] Thus, ethnicity and geographical location influence the incidence of KC directly. A summary of previous population-based studies of KC is presented in table 2.

Previous studies have shown that in the Asian population, KC is more prevalent, has an earlier onset and has a severe disease progression.[17 33] The global consensus of KC and ectatic diseases,[34] consisting of cornea subspecialists worldwide, affirms that Asians and Arabians have a higher risk of ectasia. However, the information supporting this is not entirely conclusive, as nationwide

studies also support a reduced risk of KC in the Asian population compared with Caucasians.[35] The difference may be explained by KC being more prevalent in some specific Asian populations, including Indians, Pakistanis and Saudi Arabians,[17 33] but not in East Asian people.[36 37]

Ethnographic characterisation of a South American country is complex due to its rich demographic history. Homburger et al,[38] with ancestry-specific principal component analysis, determined that the Colombian population derives primarily from European descendants. DANE (data set) reported a prevalence of 87.58% of white race, followed by 9.34% Afro-American and 4.4% Amerindian.[39] The low influence of Asian race supports a lower risk of KC in Latinos (Hispanic heritage) than the Asian population.[17 33]

KC can affect both genders and sex predilection is not apparent. Some studies report male predominance,[33 40] while others report female predominance[41] or no predilection.[28 36] In our study, males represent 65.8% of all cases, with higher incidence rates in almost all age groups, excluding 5–9, 65–69 and >80 years. Therefore, we conclude that KC in Colombia has a male predilection.

Interestingly, KC appears earlier in males (20–24 years) than in females (25–29 years). This finding has also been

Table 2 Previous population-based studies reporting the incidence rate of keratoconus

Author	Country	Publication date	Studied duration (years)	Design	Population	N	Incidence* (95% CI)
Kennedy et al[28]	USA	1986	48	Retrospective	All ages, inhabitants of Olmsted County (number not specified)	64	2 (not reported)
Ziaei et al[31]	Iran	2012	1	Prospective	990818, all ages, inhabitants of Yazd Province	221	22.3 (19.5 to 25.4)
Godefrooij et al[46]	The Netherlands	2017	4	Retrospective	1635517 inhabitants between 10 and 40 years	218	13.3 (11.6 to 15.2)
Hwang et al[36]	South Korea	2018	6	Retrospective	47990761, all ages, inhabitants	17931	5.56 (5.47 to 5.66)
Nielsen et al[16]	Denmark	2007	11	Retrospective	5326404, all ages, Danish population	–	1.3 (1.0 to 1.7)
Bak-Nielsen et al[32]	Denmark	2019	5	Retrospective	5707251, all ages, Danish population	1008	3.6 (not reported)
Cozma et al[17]	England	2005	4	Retrospective	470000, all ages, inhabitants of Bradford	197	Asian, 32.3 (not reported) White, 3.5 (not reported)
Mejia-Salgado et al (present study)	Colombia	2022	5	Retrospective	50372424 Colombian population	18419	10.36 (10.08 to 10.64)

*Per 100 000 person-years.

reported in other studies[40 42]; the question of whether males are at greater risk of aggressive disease requiring corneal transplantation has been raised. In a longitudinal, prospective, 9-year follow-up study, Edwards et al[43] found that males presented for corneal grafting earlier than females, but the overall percentage figures for surgery are similar for both males and females. Therefore, these results do not support the idea that males have a more aggressive disease than females, only that the disease is present more often and earlier in males than in females. Gender differences in the incidence of KC may be attributable to hormonal and biological changes.[44 45] However, more studies are needed to clarify these differences.

Regarding the age distribution of KC, the diagnosis incidence rate is higher in young people. Similarly, Godefrooij et al[46] considered patients between 10 and 40 years as the relevant age category for newly diagnosed KC. The peak incidence rate of KC in the general population was at 20–24 years, 28.59 (95% CI 28.50 to 28.67) cases per 100 000 inhabitants. Although corneal weakness in young ages has been the primary hypothesis of KC pathogenetic factors,[47] Kotecha et al[48] reported a negative correlation between corneal viscoelastic properties and advancing age due to an increase in cross-linkage of collagen fibrils within the cornea. These may suggest that the cornea gets stiffer with age and that KC progression significantly decreases.

Another interesting finding in our study is the onset of KC at a very young age (from 0 to 9 years of age) in 404 children; this has also been reported by Hwang et al[36] in a population-based study in South Korea between 2009 and 2014. KC commonly develops at or after teenage; however, paediatric KC can occur associated with systemic syndromes or diseases such as atopy, Down's syndrome and connective tissue disorders.[12 49] Unfortunately, we cannot assess if these cases were associated with some previously mentioned systemic diseases. In addition, although we do not know precisely how KC was diagnosed in these children, we used only confirmed new cases, so for them to be registered in SISPRO (data set) they must be diagnosed by an ophthalmology specialist using corneal topography, considering the years studied were when corneal topography use was widespread.

The strengths of this study include the large sample size. Ours is the first population-based study in Latin America to use a database from the national health registry and the most extensive study that focuses on the incidence and sociodemographic characterisation of KC worldwide. Previous studies based on SISPRO (data set) have determined the concordance rate of the database using ICD-10 with medical records to be up to 83.4%,[26 27] which means the results reported in this study offer valuable insight into the epidemiology of KC in Colombia.

Limitations

Our study has some limitations we have thoughtfully considered and addressed. First, we may have underestimated the incidence of KC since asymptomatic patients who did not use the healthcare service cannot be included. Nevertheless, we believe that we captured the most symptomatic patients considering that the Colombian health system has one of the most prominent coverages in Latin America, encompassing 50 million inhabitants, representing 97.78% of the population in 2020.[24] Underestimation and overestimation are frequent in population-based studies; nevertheless, they provide valuable information as they include a large population sample.[50]

Another limitation innate to nationwide, population-based study is the absence of uniform criteria for KC registration, as individual ophthalmologists register each case. Also, SISPRO (data set) does not allow us to see patients' identification number and thus we cannot ensure that there are no repeated patients and that duplicates are not found in cases. Nevertheless, we only include confirmed new cases to establish the incidence, avoiding most duplicates. Furthermore, we presume that an ophthalmologist confirmed each case using the gold standard criteria for the time (corneal topography), considering the years included.

CONCLUSION

This population-based study of the entire Colombian population highlights the age-specific and sex-specific incidence of KC. We found similar demographic patterns to the previous literature reports; we found an incidence of 10.36 (95% CI 10.08 to 10.64) cases per 100 000 inhabitants, an expected incidence for a Latin American population (Hispanic heritage) compared with Asian populations which report higher incidences. The incidence peaks of KC were in the early 20s in males and in the late 20s in females. The overall male to female ratio with regard to incidence was 1.60. Finally, we report the impact of the COVID-19 pandemic on the diagnosis of KC, which is essential to consider as delays in its diagnosis and treatment can lead to blindness and requirement for surgical procedures. This study provides valuable information on the epidemiology of KC in Colombia that is useful for developing policies for diagnosis, prevention and treatment of the disease.

Author affiliations
[1]Neuroscience Research Group (NEUROS), NeuroVitae Center. Institute of Translational Medicine (IMT). Escuela de Medicina y Ciencias de la Salud. Universidad del Rosario, Bogotá, Colombia
[2]Ophthalmology Interest Group, Neuroscience (NEUROS) Research Group, Neurovitae Research Center, Institute of Translational Medicine (IMT), Universidad Del Rosario Escuela de Medicina y Ciencias de la Salud, Bogotá, Colombia
[3]Grupo de Investigación Clínica, Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario, Bogotá, Colombia

Twitter William Rojas-Carabali @warc97

Contributors GM-S, CC-G, WR-C, LFP-P, LZ-P, DP: conception and design of the study, data acquisition and analysis, interpretation of data, drafting the article, revising it critically for important intellectual content, and final approval of the version to be submitted. A-d-l-T: thematic authority, conception, design of the study, data acquisition and analysis, interpretation of data, drafting the article, revising it critically for important intellectual content, and final approval of the

version to be submitted. Finally, Ad-I-T accepts full responsibility for the work and the conduct of the study, has access to the data, and controls the decision to publish.

ORCID iD
Alejandra de-la-Torre http://orcid.org/0000-0003-0684-1989

REFERENCES
1 Vazirani J, Basu S. Keratoconus: current perspectives. *Clin Ophthalmol* 2013;7:2019–30.
2 Khor W-B, Wei RH, Lim L, *et al*. Keratoconus in Asians: demographics, clinical characteristics and visual function in a hospital-based population. *Clin Exp Ophthalmol* 2011;39:299–307.
3 Santodomingo-Rubido J, Carracedo G, Suzaki A, *et al*. Keratoconus: an updated review. *Cont Lens Anterior Eye* 2022;45:101559.
4 McGhee CNJ. 2008 sir norman mcalister gregg lecture: 150 years of practical observations on the conical cornea--what have we learned? *Clin Exp Ophthalmol* 2009;37:160–76.
5 Gordon MO, Steger-May K, Szczotka-Flynn L, *et al*. Baseline factors predictive of incident penetrating keratoplasty in keratoconus. *Am J Ophthalmol* 2006;142:923–30.
6 Raiskup F, Lenk J, Herber R, *et al*. Keratokonus: heutige therapieoptionen. *Klin Monatsbl Augenheilkd* 2018;235:1148–58.
7 Hashemi H, Heydarian S, Hooshmand E, *et al*. The prevalence and risk factors for keratoconus: a systematic review and meta-analysis. *Cornea* 2020;39:263–70.
8 Assiri AA, Yousuf BI, Quantock AJ, *et al*. Incidence and severity of keratoconus in Asir province, Saudi Arabia. *Br J Ophthalmol* 2005;89:1403–6.
9 Althomali TA, Al-Qurashi IM, Al-Thagafi SM, *et al*. Prevalence of keratoconus among patients seeking laser vision correction in Taif area of Saudi Arabia. *Saudi J Ophthalmol* 2018;32:114–8.
10 Millodot M, Shneor E, Albou S, *et al*. Prevalence and associated factors of keratoconus in jerusalem: a cross-sectional study. *Ophthalmic Epidemiol* 2011;18:91–7.

11 Jurkiewicz T, Marty A-S. Correlation between keratoconus and pollution. *Ophthalmic Epidemiol* 2021;28:495–501.
12 Gordon-Shaag A, Millodot M, Shneor E, *et al*. The genetic and environmental factors for keratoconus. *Biomed Res Int* 2015;2015:795738.
13 Sahebjada S, Chan E, Xie J, *et al*. Risk factors and association with severity of keratoconus: the australian study of keratoconus. *Int Ophthalmol* 2021;41:891–9.
14 Mou Y, Qin Q, Huang X, *et al*. Risk factors and severity of keratoconus on the east coast of China. *Int Ophthalmol* 2022;42:2133–40.
15 Belin MW, Villavicencio OF, Ambrósio RR. Tomographic parameters for the detection of keratoconus: suggestions for screening and treatment parameters. *Eye Contact Lens* 2014;40:326–30.
16 Nielsen K, Hjortdal J, Aagaard Nohr E, *et al*. Incidence and prevalence of keratoconus in Denmark. *Acta Ophthalmol Scand* 2007;85:890–2.
17 Cozma I, Atherley C, James NJ. Influence of ethnic origin on the incidence of keratoconus and associated atopic disease in Asian and white patients. *Eye (Lond)* 2005;19:924–5.
18 Rafati S, Hashemi H, Nabovati P, *et al*. Demographic profile, clinical, and topographic characteristics of keratoconus patients attending at a tertiary eye center. *J Curr Ophthalmol* 2019;31:268–74.
19 Abass A, Clamp J, Bao F, *et al*. Non-orthogonal corneal astigmatism among normal and keratoconic Brazilian and Chinese populations. *Current Eye Research* 2018;43:717–24.
20 Valdez-García JE, Sepúlveda R, Salazar-Martínez JJ, *et al*. Prevalence of keratoconus in an adolescent population. *Revista Mexicana de Oftalmología* 2014;88:95–8.
21 Barraquer-Coll C, Barrera R, Molano N. Prevalencia de pacientes con queratocono en la Clínica Barraquer en Bogotá, Colombia. *Rev Soc Colomb Oftalmol* 2020;53:17–23.
22 Minsalud. Indicadores sistema integrado de información de la protección social (SISPRO). 2022. Available: https://www.sispro.gov.co/Pages/Home.aspx [Accessed 18 Apr 2022].
23 Minsalud. Registro individual de prestación de servicios de salud (RIPS). 2022. Available: https://www.minsalud.gov.co/proteccionsocial/Paginas/rips.aspx [Accessed 18 Apr 2022].
24 Minsalud. Cifras de aseguramiento en salud departamento administrativo nacional de estadistica (DANE) 2022. 2022. Available: https://www.minsalud.gov.co/proteccionsocial/Paginas/cifras-aseguramiento-salud.aspx [Accessed 18 Apr 2022].
25 Amaya J, Beltrán A, Chavarro D, *et al*. Estudio de disponibilidad y distribución de la oferta de médicos especialistas, en servicios de alta y mediana complejidad en Colombia. 2013. Available: https://www.minsalud.gov.co/salud/Documents/Observatorio%20Talento%20Humano%20en%20Salud/DisponibilidadDistribuciónMdEspecialistasCendex.pdf [Accessed 18 Apr 2022].
26 Uriben J, Gonzalez I, Cardenas D, *et al*. Lineamiento Técnico para el registro y envío de los datos del registro individual de prestaciones de salud - RIPS, desde las instituciones prestadoras de servicios de Salud a las EAPB. 2019. Available: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/OT/Lineamientos-Tecnicos-para-IPS.pdf [Accessed 18 Apr 2022].
27 Akhter M, Toy B. Big data-based epidemiology of uveitis and related intraocular inflammation. *Asia Pac J Ophthalmol (Phila)* 2021;10:60–2.
28 Kennedy RH, Bourne WM, Dyer JA. A 48-year clinical and epidemiologic study of keratoconus. *Am J Ophthalmol* 1986;101:267–73.
29 Jonas JB, Nangia V, Matin A, *et al*. Prevalence and associations of keratoconus in rural Maharashtra in central India: the central India eye and medical study. *American Journal of Ophthalmology* 2009;148:760–5.
30 Özalp O, Atalay E, Yıldırım N. Prevalence and risk factors for keratoconus in a university-based population in Turkey. *J Cataract Refract Surg* 2021;47:1524–9.
31 Ziaei H, Jafarinasab MR, Javadi MA, *et al*. Epidemiology of keratoconus in an Iranian population. *Cornea* 2012;31:1044–7.
32 Bak-Nielsen S, Ramlau-Hansen CH, Ivarsen A, *et al*. Incidence and prevalence of keratoconus in Denmark - an update. *Acta Ophthalmol* 2019;97:752–5.
33 Pearson AR, Soneji B, Sarvananthan N, *et al*. Does ethnic origin influence the incidence or severity of keratoconus? *Eye* 2000;14:625–8.
34 Gomes JAP, Tan D, Rapuano CJ, *et al*. Global consensus on keratoconus and ectatic diseases. *Cornea* 2015;34:359–69.
35 Woodward MA, Blachley TS, Stein JD. The association between sociodemographic factors, common systemic diseases, and keratoconus. *Ophthalmology* 2016;123:457–465.

36 Hwang S, Lim DH, Chung T-Y. Prevalence and incidence of keratoconus in south korea: a nationwide population-based study. *Am J Ophthalmol* 2018;192:56–64.

37 Ota R, Fujiki K, Nakayasu K. [Estimation of patient visit rate and incidence of keratoconus in the 23 wards of tokyo]. *Nippon Ganka Gakkai Zasshi* 2002;106:365–72.

38 Homburger JR, Moreno-Estrada A, Gignoux CR, *et al*. Genomic insights into the ancestry and demographic history of South America. *PLoS Genet* 2015;11:e1005602.

39 Departamento Adminitrativo Nacional de Estadistica (DANE). Grupos étnicos - información técnica. n.d. Available: https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/grupos-etnicos/informacion-tecnica

40 Owens H, Gamble G. A profile of keratoconus in New Zealand. *Cornea* 2003;22:122–5.

41 Krachmer JH, Feder RS, Belin MW. Keratoconus and related noninflammatory corneal thinning disorders. *Surv Ophthalmol* 1984;28:293–322.

42 Fink BA, Wagner H, Steger-May K, *et al*. Differences in keratoconus as a function of gender. *Am J Ophthalmol* 2005;140:459:459–68.:.

43 Edwards M, Clover GM, Brookes N, *et al*. Indications for corneal transplantation in New Zealand: 1991-1999. *Cornea* 2002;21:152–5.

44 Harding JJ, Rixon KC, Marriott FHC. Men have heavier lenses than women of the same age. *Exp Eye Res* 1977;25:651.

45 Goto T, Klyce SD, Zheng X, *et al*. Gender- and age-related differences in corneal topography. *Cornea* 2001;20:270–6.

46 Godefrooij DA, de Wit GA, Uiterwaal CS, *et al*. Age-specific incidence and prevalence of keratoconus: a nationwide registration study. *Am J Ophthalmol* 2017;175:169–72.

47 Malik N, Moss S, Ahmed N, *et al*. Ageing of the human corneal stroma: structural and biochemical changes. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1992;1138:222–8.

48 Kotecha A, Elsheikh A, Roberts CR, *et al*. Corneal thickness- and age-related biomechanical properties of the cornea measured with the ocular response analyzer. *Invest Ophthalmol Vis Sci* 2006;47:5337–47.

49 Davidson AE, Hayes S, Hardcastle AJ, *et al*. The pathogenesis of keratoconus. *Eye (Lond)* 2014;28:189–95.

50 Thygesen LC, Ersbøll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol* 2014;29:551–8.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open Ophth

1   **Supplementary material 1: statical analysis**

2

3   In this work, we use a statistical model for spatial data to make a prediction about what would have happened

4   in the year 2020 on the incidence of the disease, this is done since the results are greatly affected by the

5   pandemic. The statistical models for spatial data are divided by Cressie [1] into two broad classes:

6   geostatistical models with continuous spatial support and models in a lattice, also called area models [2], where

7   the data occur in a (possibly irregular) grid, with an enumerable set of vertices or locations . The two most

8   common area models are conditional autoregressive (CAR) models and simultaneous autoregressive (SAR)

9   models. These autoregressive models are used in many fields, including mapping disease rates [3], agriculture

10   [4], econometrics [5], ecology [6] and image analysis [7]. In this paper we will focus on CAR models. CAR

11   models are an example of the Gaussian Markov random fields [8] and the popular nested Laplace approach

12   integrated methods [9].

13

14   The basis of these models is the Gaussian Markov Fields. Random fields are multivariate distributions that are

15   generally used to describe the spatial association between variables X. A Markov random field extends the

16   Markov chain concept to a spatial context and assumes that such a joint distribution of X satisfies:

17
$$f(X_i|\boldsymbol{X}_{-i}) = f(X_i|\boldsymbol{X}_{j\sim i}),$$

18   where, $\boldsymbol{X}_{j\sim i}$ is the vector formed by all the components of X that are neighbors of i. A Gaussian Random

19   Markov Field (GMRF) is a Markov field where the random vector distribution (finite-dimensional) is a normal

20   or Gaussian distribution satisfying the conditional independence assumptions.

21

22   An n-dimensional random vector $y_{n\times 1} = (y_1, y_2, ..., y_n)^T$, $n < \infty$ has a n$-$variable distribution with mean vector $\mu_{n\times 1}$

23   and covariance matrix $\Sigma_{n\times n}$, and its probability density function (fdp) assumes the as follows:

24

25
$$f_{\boldsymbol{y}}(\boldsymbol{y}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\}.$$

26

27   This distribution will be denoted by $\boldsymbol{y} \sim N(\boldsymbol{\mu}, \Sigma)$ where μ and Σ only have:

28

$$\mu_i = E(y_i)$$

and $\Sigma_{ij} = Cov(y_i, y_j), \Sigma_{ii} = Var(y_i) \text{ and } Corr(y_i, y_j) = \Sigma_{ij}(\Sigma_{ii}\Sigma_{jj})^{-1/2}.$

31

To build a GMRF we consider a graph G = (V, E) with n vertices where each vertex represents one of the components of the vector y = (y₁,y₂,...,yₙ) and edges connect nodes that have some sort of association. A GMRF assumes that $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and that the edges of the graph connect nodes i and j if and only if $y_i \perp y_j | y_{-ij}$, that is, if y_i is independent of y_j, given the components of y except y_i and y_j. In a GMRF, the covariance matrix brings information about the connections between the nodes through the precision matrix $\Sigma^{-1} = Q$ which is a symmetric matrix and positive definite.

So, a random vector y = (y₁,y₂,...,yₙ)ᵀ ∈ Rⁿ is called GMRF corresponding to a graph G = (V,E) with mean μ and precision matrix Q > 0, if and only if the pdf of y has the following form:

$$\pi(\boldsymbol{y}) = (2\pi)^{-n/2}|\boldsymbol{Q}|^{1/2}\exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{Q}(\boldsymbol{y}-\boldsymbol{\mu})\right),$$

where the array Q satisfies the following condition:

$$Q_{ij} \neq 0 \Leftrightarrow \{i, j\} \in \mathcal{E}, \forall i \neq j.$$

If Q is a completely dense matrix, then G is fully connected, that is, the vertex is connected to all other vertices in the graph. Let's focus on the case where Q is sparse. All results valid for the normal distribution will also be valid for a GMRF. A detailed discussion of GMRF can be found in Rue and Held [8].

An example of a GMRF is the conditional autoregressive model or CAR model, in this case we consider a geographic region that is partitioned into n subregions indexed by integers 1,2,...,n and assume that this

2

53 collection of sub-regions has a neighborhood system $\{V_i : i \; 1,...,n\}$, where V_i denotes the collection of sub-

54 regions that, in a well-defined sense, are neighbors of the subregion i. In geographical terms,

55

56 $V_i = \{j :$ the subregions i and j share a boundary$\}$, to $i \in \{1,2,...,n\}$,

57

58 The neighborhood system is a key point in autoregressive or CAR models that are commonly used in spatial

59 statistics, the graphs that support the construction of the GMRF will be those that express these neighborhood

60 structures. In this context, the edges E in the graph G = (V, E), represent the connections in the geographic

61 structure and, consequently, define the neighbors that are used to model spatial dependence. The components

62 of the vector y are nodes of the graph.

63

64 Let $y_1,...,y_n$ be the observations made in the areas 1,...,n. Let us denote by $j \sim i$ that node j is a neighbor of node

65 i. The term conditional, in the CAR model is used because each element of the random process is conditionally

66 specified in the values of neighboring nodes, the CAR model assumes that the complete conditional

67 distributions are normal distributions. Then, we assuming that,

68

$$y_i | y_{-i} \sim N(\mu_i + \rho_{\mathcal{G}} \overline{(y - \mu)}_i, \frac{\sigma_{\mathcal{G}}^2}{d_i^{\mathcal{G}}})$$

69 ,

70

71 where $\sigma_{\mathcal{G}}^2 / d_i^{\mathcal{G}}$ is the conditional variance of $y_i | y_{-i}$, $\rho_{\mathcal{G}}$ is a proportionality constant, $d_i^{\mathcal{G}}$ is the number of

72 neighbors of node i in the graph G, the average of the neighbors of node i is the:

$$\overline{(y - \mu)}_i = \sum_{\mathcal{E}^{\mathcal{G}}} (d_i^{\mathcal{G}})^{-1} (y_j - \mu_j)$$

73

74 and,

$$\mathcal{E}^{\mathcal{G}} = \{(i,j) \in E(\mathcal{G}) : j \sim i\}$$

75

76 is the set of edges that belong to the graph G. Consider the adjacency matrix $A_{\mathcal{G}} = (a_{ij})$ such that $a_{ii} = 0$, $a_{ij} = 1$

77 if i and j $a_{ij} = 0$ if i $6=\sim$ j and $M_{\mathcal{G}} = diag\{d_1^{\mathcal{G}}, d_2^{\mathcal{G}}, \ldots, d_n^{\mathcal{G}}\}$.

3

78

79   Besag uses Brook's lemma [2, 10] and shows that when the matrix $(M_{\mathcal{G}} - \rho_{\mathcal{G}} A_{\mathcal{G}})^{-1}$ is positive definite and

80   symmetric the joint distribution for y is:

$$\boldsymbol{y} \sim N(\boldsymbol{\mu}, (\Sigma_{CAR}^{\mathcal{G}})^{-1}),$$

81

82

83   where $(\Sigma_{CAR}^{\mathcal{G}})^{-1} = \sigma_{\mathcal{G}}^2 (M_{\mathcal{G}} - \rho_{\mathcal{G}} A_{\mathcal{G}})^{-1}$. For the covariance matrix to be positive definite, it

84   is necessary that $\rho_{\mathcal{G}} < \frac{1}{\lambda_1}$ where λ_1 is the smallest eigenvalue of the matrix $M_{\mathcal{G}}^{-1/2} A_{\mathcal{G}} M_{\mathcal{G}}^{-1/2}$ Banerjee et al

85   [2].

86

87   In conclusion, the CAR model approach visualizes the geographic domain as an undirected graph with a vertex

88   in each region and an edge between two vertices if the corresponding regions share a geographic edge. This

89   creates well-defined neighbors for each region, which are used to define the joint or conditional distribution.

90   The distribution will be the multivariate normal distribution. All analysis of the CAR model is concentrated

91   on the covariance matrix Σ, which is defined by the graph of the geographic domain and the parameter ρ.

92

93   **REFERENCES:**

94   1. Cressie NAC. Statistics for spatial data. Rev. ed. New York: Wiley; 1993.

95   2. Banerjee S, Carlin BP, Gelfand AE, Banerjee S. Hierarchical Modeling and Analysis for Spatial Data. 0

96   edition. Chapman and Hall/CRC; 2003.

97   3. Elliott P, Wartenberg D. Spatial Epidemiology: Current Approaches and Future Challenges. Environmental

98   Health Perspectives. 2004;112:998–1006.

99   4. Besag J, Higdon D. Bayesian analysis of agricultural field experiments. J Royal Statistical Soc B.

100   1999;61:691–746.

101   5. LeSage J, Pace RK. Introduction to Spatial Econometrics. 0 edition. Chapman and Hall/CRC; 2009.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open Ophth

102   6. Arslan O, Akyürek Ö. Spatial Modelling of Air Pollution from PM10 and SO2 concentrations during Winter

103   Season in Marmara Region (2013-2014). International Journal of Environment and Geoinformatics. 2018;5:1–

104   16.

105   7. Besag J. On the Statistical Analysis of Dirty Pictures. Journal of the Royal Statistical Society Series B

106   (Methodological). 1986;48:259–302.

107   8. Rue H, Held L. Gaussian Markov Random Fields. 0 edition. Chapman and Hall/CRC; 2005.

108   9. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using

109   integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series B (Statistical

110   Methodology). 2009;71:319–92.

111   10. Besag J. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical

112   Society Series B (Methodological). 1974;36:192–236.

113

1 **Supplementary material 2: Distribution by department of the incident keratoconus patients**

2 **based on SISPRO** [1]**.**

Departament	Number of patients with keratoconus	% of keratoconus by department
Bogotá, D.C.	8,959	48.63
Antioquia	2,587	14.04
Dep_nodef	1,771	9.61
Cundinamarca	1,912	10.38
Valle del Cauca	609	3.3
Nariño	622	3.37
Boyacá	468	2.54
Caldas	209	1.13
Cauca	290	1.57
Santander	211	1.14
Quindio	43	0.23
Risaralda	71	0.38
Tolima	112	0.60
Atlántico	81	0.43
Huila	76	0.41
Meta	66	0.35
Córdoba	65	0.35
Norte de Santander	48	0.26
Sucre	30	0.16
Bolívar	33	0.17
Cesar	42	0.22

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open Ophth

Putumayo	27	0.14
Magdalena	33	0.17
Casanare	7	0.03
Caquetá	17	0.09
Chocó	10	0.05
La Guajira	8	0.04
San Andrés, Providencia	1	0.01
Arauca	2	0.01
Guaviare	5	0.03
Vaupés	3	0.02
Amazonas	0	0,00
Guainía	1	0.01
Vichada	0	0.00
Total	18,419	100.00

3

4 **REFERENCES:**

5 1. Minsalud. Indicadores Sistema Integrado de Información de la Protección social (SISPRO). 2022.

6 https://www.sispro.gov.co/Pages/Home.aspx. Accessed 18 Apr 2022

7