medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Estimating spatial disease rates from health statistics without geographic identifiers

Javier Cortes-Ramirez [1,2,3], Juan D. Wilches-Vega [2], Ruby N. Michael [4], Vishal Singh [5], Olga M. Paris-Pineda [2]

1. Centre for Data Science, Queensland University of Technology, Brisbane, Australia
2. Faculty of Health, University of Santander, Cúcuta, Colombia
3. Children's Health and Environment Program, University of Queensland, Brisbane, Australia
4. Cities Research Institute, Griffith University, Brisbane, Australia.
5. School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia

Corresponding author:
Dr Javier Cortes-Ramirez.
O Block D Wing Room D722. Ring Road, Kelvin Grove Campus
Queensland University of Technology, Victoria Park Road. Kelvin Grove QLD 4059 Australia
P: +61 (07) 3138 5278. E: javier.cortesramirez@qut.edu.au

Abstract
Effective analysis to support decision-making in public health requires adequate data that can be linked to the sociodemographic characteristics and distribution of communities that would be recipients of health programs. Although official statistics are an important source of data to support decisions on public health strategies, health statistics are of limited use if they do not include categorisations related to the spatial distribution of the population. This study introduces a method to use the frequency of disease reported by health care providers (HCP) to calculate disease rates in geographical areas in urban settings. Specifically, this study uses statistics on acute respiratory infections (ARI) to calculate the rate of these diseases at the census district level in Cúcuta, a major city in Colombia. A Geographic Information System was used to establish the geographical area of influence (spatial scope) of each HCP, according to the distribution of the basic census geographical areas; the Urban Sections (USEC), in Cúcuta. Three levels of increasing spatial aggregation were established considering the characteristics of the population receiving health care in primary, intermediate and tertiary public health services to establish the spatial scope of each HCP. The cases of ARI per USEC were calculated according to the proportion of the population of each USEC in each of the three levels. The spatial rate of ARI per USEC and the hotspots of higher risk of ARI were calculated using an Empirical Bayes method using a geostatistical software.

There were 97 HCP, of which, 31 provided health services in USEC in level 1; 20 provided health services in USEC in level 2; and 47 provided health services in USEC in level 3. A higher spatial rate of ARI was found in USEC in central south; central west; north and northwest; northeast; central east; and central regions, compared to the whole of Cúcuta. Hotspots of higher risk were identified in two clusters of USEC in central south and west Cúcuta and three isolated USEC in central and northwest Cúcuta.

Health indicators at the census district geographical level could be calculated using basic statistics that do not include geographic identifiers by considering the characteristics of the HCP and the population receiving their services. The methodology of this study can be applied to other socioeconomic contexts where geographic identifiers are not attached to public health statistics to create better public health indicators and support potential health promotion and disease prevention strategies.

Key words: Census district boundaries, spatial epidemiology, Local Indicators of Spatial Autocorrelation -LISA, morbidity, Cúcuta –North of Santander;

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Effective decision making in public health is underpinned by using the best available medical evidence and research on risk factors determinants of disease. An important condition to conduct effective analysis in research is the availability and quality of data that can be used in various quantitative methodological frameworks.[1] Statistics and health reports are especially important for epidemiological studies to identify the association of risk factors with health outcomes in the general population and support potential prevention strategies. However, statistical reports can be inadequate to provide the best evidence unless they are designed for specific analyses for at-risk populations.[2] Decisions related to diagnostic and therapeutic actions are usually based on high-quality research studies such as clinical trials. In contrast, public health decisions, especially by local authorities, are often based on the review of basic statistics with limited academic rigour.[3] The lack of high-quality data can favour the role that politics and political ideology have on shaping local governments initiatives in public health which can impact effective evidence-based decision making.[4] This can be accentuated in low and middle-income countries where the public health sector has limited access to resources for population health research.[5]

Public health departments usually collect statistics from health providers such as hospitals, general practices and clinics and health insurance providers to monitor public health. The utility of these data depends on the inclusion of individual identifiers or categories such as socioeconomic status or residential areas. Public health data linked to small geographical areas such as counties or local government areas can be used to identify the distribution of diseases or mortality in cities or larger regions.[6] In addition, census information for geographical areas can be linked to the public health data to calculate proportional rates of cases by population and the association of these proportions with other indicators such as socioeconomic and education indexes. This comes with challenges, in that basic health statistics may include unformatted and incomplete data and a lack of information at the individual level. This paucity of information is more evident in resource-constrained countries with up to 152 low-income and middle-income countries having no data or poor-quality health data.[7]

The North of Santander region in Colombia has an important burden of disease due to respiratory infections with the capital city, Cúcuta, accounting for more than 80% of the respiratory morbidity in the region.[8,9] Cúcuta has a population of 711,715 inhabitants in an area of 1,119 Km^2 that access a complex network of public and private health services of primary, intermediate and tertiary healthcare. The population is distributed in 460 geographical areas defined by census district boundaries called urban sections (USEC). Despite the large number of USEC and the socioeconomic diversity between these districts, the statistics of acute respiratory infections (ARI) can be obtained only at the health centres and hospital level. The usability of the statistics of ARI in Cúcuta for analyses and assessment to support decision-making in public health is very limited because it is not linked to the distribution of the population across the city. This prevents the calculation of useful health indicators such as the rate of ARI per USEC.

Although the statistics of ARI in Cúcuta are not provided at a geographical level, the geographical distribution of the health care providers can still be used to make estimations of the number of cases occurring within the USEC. Spatial statistics can be used to estimate the distribution of observations in well-defined geographic settings using assumptions based on known characteristics of individuals or other agents to interpolate missing observations and compensate for the lack of data.[10,11] A similar approach can be used to estimate the cases of ARI in Cúcuta, considering the spatial distribution of the health care providers and the population in the USEC. The aim of this study is to use statistics reported by the health care providers (HCP) to increase the quality of the morbidity data in Cúcuta to support decision-making in public health. The first objective is to design a method to calculate and map the spatial rate of ARI per USEC in Cúcuta, using statistics of cases reported at the HCP level. The second objective is to identify the hotspots of higher risk of ARI. This will be achieved using geographic information system (GIS) techniques to establish the geographical area of

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

influence for each HCP in Cúcuta, to estimate the ARI cases per USEC, and the use of geostatistical software to identify the spatial hotspots of higher risk.

Methods

Data on the number of medical consultations with a diagnosis of ARI in the period 01-January-2018 to 31-December-2018, were provided by the Public Health Department categorised by HCP. As these data were not linked to geographical areas directly, the number of ARI consultations per geographical area in Cúcuta was estimated by categorising HCP into USEC —that are the basic geographic units established by the National Department of Statistics (Figure 1). This was done by defining spatial zones around each HCP according to the spatial extent of areas receiving their services (i.e., spatial scope). Three types of HCP were identified: primary HCP that provide primary health services to population located in specific geographical areas; intermediate HCP that provide intermediate health care to people in larger areas of the city; and tertiary hospitals and nonselective private HCP and general practitioners (GPs) that provide health services to people from any area in the city. To establish the spatial zones around each HCP, the spatial distribution of the primary and intermediate public HCP was considered because these are distributed to provide services to populations in specific geographic sectors across the city (Figure 1).

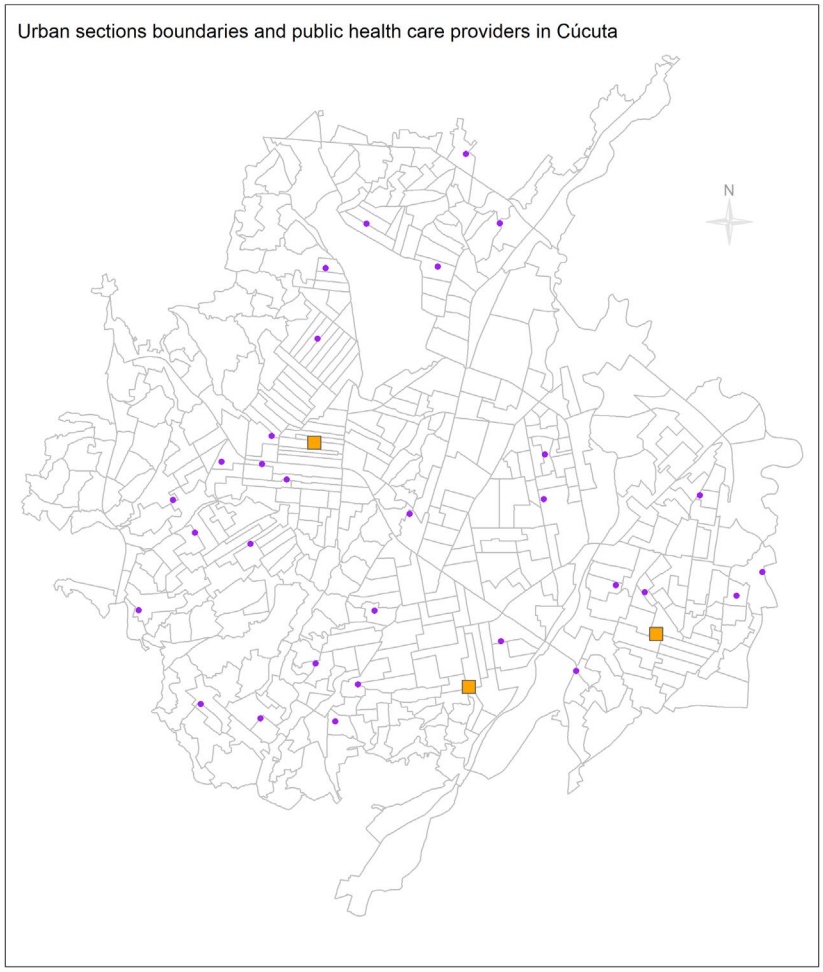

Figure 1. Spatial distribution of the Cúcuta Urban Sections and the primary (dots) and intermediate (squares) public health care providers.

Three levels of increasing spatial aggregation were established based on the spatial scope of the public HCP: Level 1 zones were the most tightly defined and included the USEC within the spatial scope of each single public primary HCP. Level 2 zones were defined to include the USEC in the spatial scope of the three public intermediate HCP. The Level 3 zone included all Cúcuta's USEC and

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

was defined as the spatial scope of the hospitals and nonselective private HCPs. The spatial scope for level 1 and 2 zones was defined with ARCmap (v.10.6) using the nearest neighbour algorithm (i.e., identifies the USEC that is geographically closest to another one). Some level 2 zones were edited manually to adjust for the USEC connectivity according to the bus routes and the Pamplonita river shape.

To estimate the number of ARI consultations in each USEC, a weight value was assigned according to the proportion of population of each USEC within each level zone. For example, the weight value of a USEC with 500 people in a level 1 zone with a total population of 5000 people is 500 ÷ 5000 = 0.1. The total ARI consultations per zone in each level were multiplied by the weight value to obtain the ARI cases per USEC in each zone and level. Each USEC would have ARI cases calculated for each of the three levels, therefore, the total ARI per USEC was calculated as the sum of USEC ARI in each level (i.e., total ARI in each USEC= ARI Level 1 + ARI Level 2 + ARI Level 3).

Once the total ARI cases were calculated for each USEC, the incidence rate of ARI per USEC were calculated as the total ARI divided by the USEC population. However, the ARI cases per USEC calculated on the assumption of an even distribution of cases per population can produce distortion of the estimated ARI cases and difficulties to compare the rates between USEC due to differences in the underlying population. To overcome these issues, the rates can be smoothed using spatial statistics techniques. The most common method to calculate spatially smoothed rates uses Bayesian principles, particularly the Empirical Bayes (EB) approach.[12] To calculate the spatial rate of ARI per USEC (ARI-sr) in Cúcuta, spatial EB rates were estimated with Geoda (v. 1.18). A similar approach can be used to assess the spatial trend ARI in the Cúcuta's USEC to identify statistically significant hotspots of higher risk of ARI.[13] The hotspots were identified using the local Moran's I test with EB rate in Geoda. Maps were drawn in Rstudio ver. 1.2 using the R-package Tmap.[14]

This study was approved by the ethics committee of the University of Santander, record VII-FT-025-UDES, 021 25/06/2019.

Results

There were 118,469 cases of ARI in Cúcuta over the study period. Of these, 38,236 (32.3%) were reported by hospitals; 43,173 (36.4%) were reported by intermediate HCP; and 37,060 (31.3%) reported by primary health centres or GPs. Table 1 shows the number of ARI cases for each spatial level of aggregation.

Table 1. Cases of acute respiratory infections per Health Care Provider and spatial level of aggregation

Type of Health Care Provider	N	ARI cases	Level 1	Level 2	Level 3
Hospitals	5	27732			27732
Intermediate health units	19	53677		14419	39258
Primary health centres and GPs	74	37060	8994	4031	24035

N: number of health care providers. ARI: Acute Respiratory Infections. GPs: General practitioners.

There were 97 HCP that included 5 hospitals that provided health care to the population in all USECs (level 3). There were 19 intermediate HCP of which; 14 provided services to the population in all USECs, and 5 provided health services to the population in USECs in level 2 zones. There were 74 Primary HCP or GPs of which; 28 provided health care to the population in USECs in the level 3 zone; 15 provided health services to the population in USECs in level 2 zones, and 31 provided health care to the population in USECs in level 1 zones (details in the appendix). Three level 2 zones were identified from the area of influence of the intermediate public HCP (Figure 2a) and 26 level 1 zones were identified from the area of influence of the public primary HCP (Figure 2b).

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

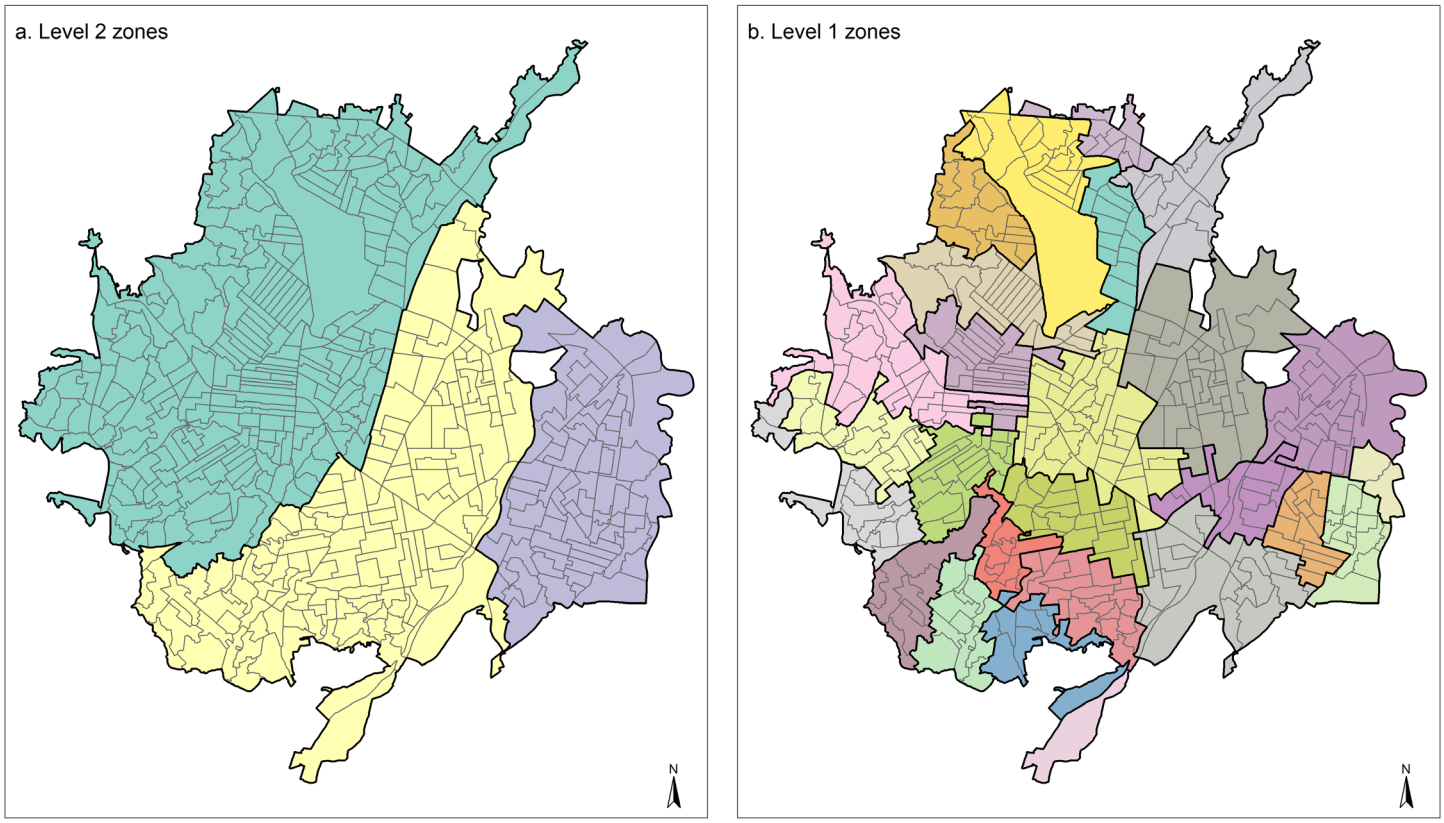

Figure 2. Spatial distribution of the geographic zones for levels of aggregation 2 (a) and 1 (b)

The spatial rate of ARI (ARI-sr) ranged from 0.16 to 3.65 per inhabitant with three USEC having more than one ARI case per inhabitant (i.e., ARI-sr >1). Figure 2.a shows the distribution of the ARI-sr in Cúcuta in 2018.

Higher spatial rates of ARI were found in clusters in central south; central west; north and northwest; northeast; central east; and central Cúcuta, compared to the whole of Cúcuta. Figure 2.b shows the location of USECs with significantly higher risk of ARI that have neighbour USEC also with significantly higher ARI-sr (i.e., statistically significant hotspots of higher risk). Two clusters of hotspots in central south and west Cúcuta and three isolated hotspots in central and northwest Cúcuta were identified.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

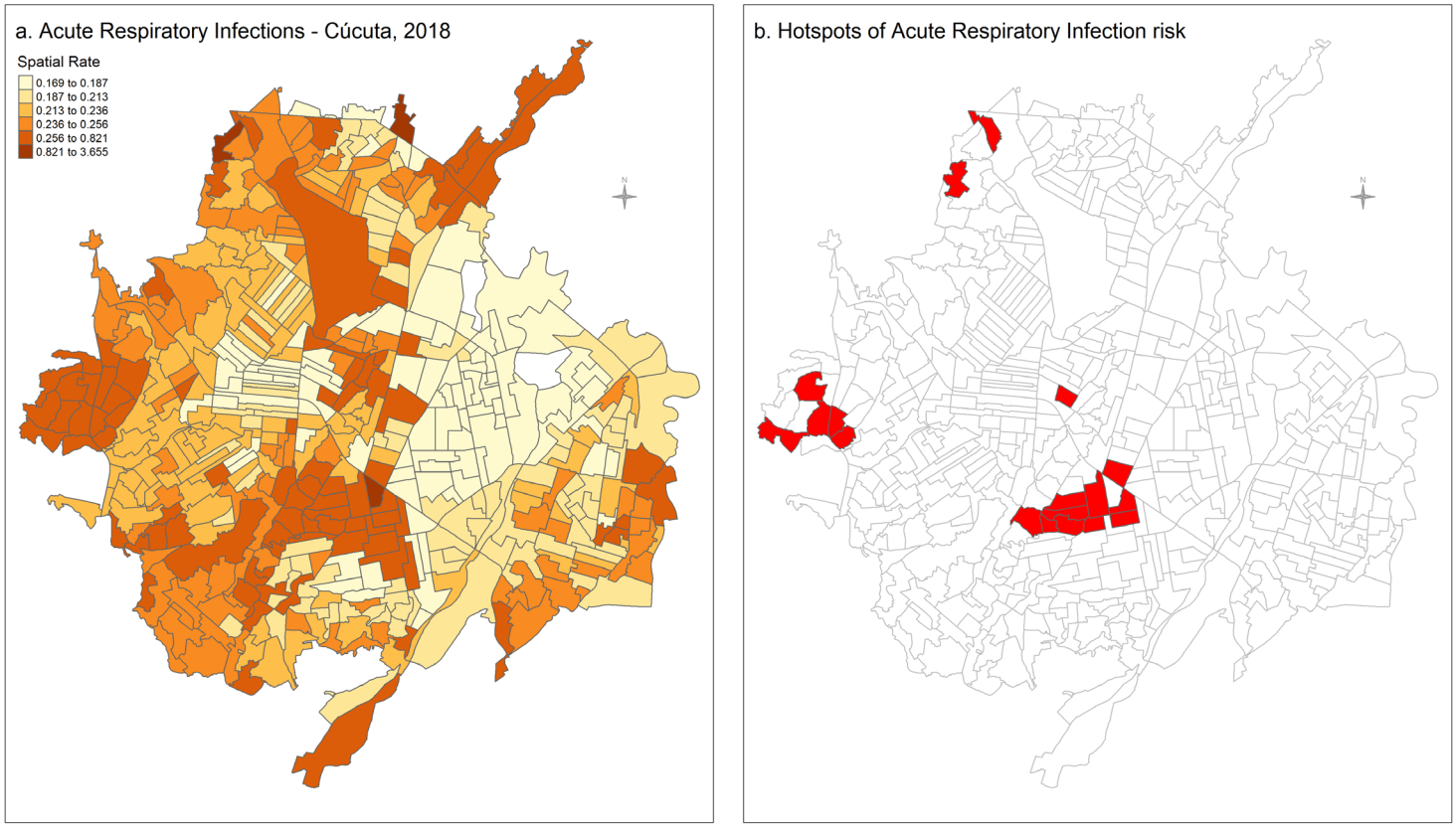

Figure 2. Spatial rates of Acute Respiratory Infections (a) and statistically significant hotspots of risk of ARI (b) in Cúcuta, 2018.

Discussion

In this study, we calculated and mapped the highest spatial rates of Acute Respiratory Infections (ARI) in multiple urban sections (USEC) in central and north regions of Cúcuta in 2018 and identified the statistically significant hotspots of higher risk of ARI. We used a GIS to identify the USEC within the area of influence of each health care provider, implementing a novel methodology to assign the number of ARI consultations to each USEC according to three geographical levels of aggregation in the city. Our method shows that health data from health care providers can be used to estimate the rate of health outcomes such as ARI, per small geographical areas, using the census district boundaries when the spatial scope of the HCP is considered. This is the first study to estimate the rate of respiratory infections in census district areas in a major city in the Colombia northeast region.

The availability of relevant and adequate data is a basic condition for public health departments to deal with many challenges including efficient health care provision and planning. The analysis of data at a geographical level is an important approach to address public health challenges with the help of increasingly used tools such as GIS and geostatistical software.[15] We used GIS techniques and spatial data science software in combination with assumptions on the Cúcuta health care institutions and population characteristics to estimate morbidity indicators at a small geographical level. Spatial analyses of health data have several advantages that can be used to support decision making including the identification of spatial autocorrelation and detection of spatial clusters.[16] The potential distortion introduced by having assumptions about the distribution of cases in geographical urban areas can be reduced by the smoothing effect that is incorporated in spatial rate calculations to increase the robustness of the estimates.[17] Whereas the health department of cities such as Cúcuta might not have data on morbidity per census districts levels, their calculation using the method used in this study can be incorporated as an instrument to support planning strategies or prioritise resources in areas of interest.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

An important advantage of having indicators at the census districts level is the possibility of assessing their relationship with other indicators such as indexes of socioeconomic development and demographics characteristics including ethnicity, gender and age groups. This can increase the utility of the morbidity rates calculated in this study to produce more specific and complex analyses. For example, correlation tests can be used to explore predictive relationships between health outcomes and some risk factors of interest.[18] These measures are often considered to better understanding the potential effect of environmental exposures on specific diseases and their relationship with public health interventions.[19,20] For example, the spatial rate of ARI can be correlated with important factors such as the index of multidimensional poverty and education levels that are measured at the USEC level in Cúcuta. Having data of multiple factors can also be used for multivariate analysis using a regression model to estimate the relative influence of one or more predictors on the ARI rates and to identify outliers or anomalies.[21] This is particularly important in studies at a geographical level because of the risk of ecological bias, where confounding factors at the group level can produce spurious associations. The use of spatial statistic techniques such as the design of mixed-effects spatial regressions using the spatial units as the random effects can reduce the risk of ecological bias.[22] The calculation of rates presented in this study can be extended to identify potentially associated risk factors by designing further multivariate analyses.

Although we used census districts to build a health indicator such as the spatial rate of ARI, this methodology can be applied to other types of data such as community participation and political or electoral involvement. Although measuring community participation to support decision-making in public health is one of the most significant challenges faced by the public health and academic sectors,[23] it is recognised that communities have a key role in the planning and organising of health care.[24] On the other hand, in some socioeconomic contexts, especially middle and low-income countries, electoral and administrative districts do not overlap or there are significant gaps between rural and urban areas, making it harder to identify patterns that align with the sociodemographic distribution of the population.[25] The use of GIS techniques to improve the resolution of health statistics used in this study is a methodological approach that can be implemented to do analyses of data in social science. This can also facilitate the implementation of other emerging tools to identify and measure socio-political determinants in public health such as System Dynamics Modelling[26] and Structural Equation Modelling.[27]

Our study had some limitations, especially regarding the data for specific individual markers such as age or gender that would have allowed the estimation of age and sex-adjusted spatial rates. Age-adjusted rates are used because crude rates of populations with different frequencies in specific age groups (i.e., age structure) can be misleading.[28] To reduce the potential effect of the age structure of the population, we calculated spatial rates that smooth the estimates for each spatial unit, removing the effect of outliers and taking into account the spatial trend or the crude rate to increase robustness.[12] However, the standardisation of rates by age and or sex should be considered in further analyses using these data.

We had access to health data of 2018 and could match these data with the 2018 census districts although the inclusion of data matching more than one census year could have allowed identifying a temporal trend, producing stronger estimates. However, this would be effective in contexts with a periodic census which is not the case in Colombia where the previous census was in 2005, and the spatial trend of the census districts is less reliable.

Conclusions

Basic health statistics can be used to estimate health indicators at the same geographical level of the census districts areas even when these data do not include geographical identifiers. This study established geographical levels of aggregation considering the spatial scope of health care providers to calculate the spatial rate of acute respiratory infections in the urban sections of Cúcuta, Colombia.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

This approach also allowed the identification of hotspot or areas of higher risk to support public health decision making. The methodology developed in this study can be used to enhance the usability of public health statistics where geographic identifiers are not included at the individual level, especially in low- and middle-income countries where restrictions due to limited resources affect the availability of adequate and quality health data.

Authors contribution: JC: conceptualisation, formal analysis and writing (original draft). JDWV: Data curation, writing (review & editing). RNM: Validation, writing (review & editing). VS: Validation, writing (review & editing). OMPP: Data curation, writing (review & editing).

Data availability statement: All data produced in the study are available upon reasonable request to the authors. The raw data provided by the Department of Health is subject to their authorisation to be published.

Ethical approval: This study was approved by the ethics committee of the University of Santander, record VII-FT-025-UDES, 021 25/06/2019.

Informed consent: N/A

Registry and registration no. of the study/trial: N/A

Conflict of interest: The authors declare no conflict of interest. This study did not receive funding.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. European Centre for Disease Prevention and Control. The use of evidence in decision-making during public health emergencies. *Stockholm: ECDC.* 2019.
2. Kneale D, Rojas-García A, Thomas J. Exploring the importance of evidence in local health and wellbeing strategies. *Journal of Public Health.* 2018;40(suppl_1):i13-i23.
3. McGill E, Egan M, Petticrew M, et al. Trading quality for relevance: non-health decision-makers' use of evidence on the social determinants of health. *BMJ Open.* 2015;5(4):e007053.
4. Kneale D, Rojas-García A, Thomas J. Obstacles and opportunities to using research evidence in local public health decision-making in England. *Health Research Policy and Systems.* 2019;17(1):61.
5. Chopra M. Inequalities in health in developing countries: Challenges for public health research. *Critical Public Health.* 2005;15(1):19-26.
6. Cortes-Ramirez J, Vilcins D, Jagals P, Soares Magalhaes RJ. Environmental and sociodemographic risk factors associated with environmentally transmitted zoonoses hospitalisations in Queensland, Australia. *One Health.* 2020;12:100206-100206.
7. Boerma JT, Stansfield SK. Health statistics now: are we making the right investments? *The Lancet.* 2007;369(9563):779-786.
8. Gobierno de Colombia. Morbilidad consulta externa. Medicina General. 2019. *Sistema de Datos Abiertos.* 2021;Retrieved from https://www.datos.gov.co/browse?Informaci%C3%B3n-de-la-Entidad_Departamento=Norte+de+Santander&Informaci%C3%B3n-de-la-Entidad_Municipio=C%C3%BAcuta&category=Salud+y+Protecci%C3%B3n+Social&page=2.
9. Instituto Nacional de Salud. Infección respiratoria aguda. Semana epidemiológica 40. *Boletin Epidemiologico Semanal.* 2019;retrieved from: https://www.ins.gov.co/buscador-eventos/BoletinEpidemiologico/2019_Boletin_epidemiologico_semana_40.pdf.
10. Griffith DA, Bennett RJ, Haining RP. Statistical Analysis of Spatial Data in the Presence of Missing Observations: A Methodological Guide and an Application to Urban Census Data. *Environment and Planning A: Economy and Space.* 1989;21(11):1511-1523.
11. Boakes EH, Fuller RA, McGowan PJK, Mace GM. Uncertainty in identifying local extinctions: the distribution of missing data and its effects on biodiversity measures. *Biology Letters.* 2016;12(3):20150824.
12. Anselin L, Lozano N, Koschinsky J. Rate transformations and smoothing. *Urbana.* 2006;51:61801.
13. Anselin L. Local Indicators of Spatial Association—LISA. *Geographical Analysis.* 1995;27(2):93-115.
14. Tennekes M. tmap: Thematic Maps in R. *Journal of Statistical Software.* 2018;84(1):1-39.
15. Chinnaswamy A, Papa A, Dezi L, Mattiacci A. Big data visualisation, geographic information systems and decision making in healthcare management. *Management Decision.* 2019.
16. Kirby RS, Delmelle E, Eberth JM. Advances in spatial epidemiology and geographic information systems. *Annals of Epidemiology.* 2017;27(1):1-9.
17. Blangiardo M, Cameletti M. *Spatial and spatio-temporal Bayesian models with R-INLA.* John Wiley & Sons; 2015.
18. Franco F, Di Napoli A. Measures of Association in Medicine and Epidemiology. *Giornale di Tecniche Nefrologiche e Dialitiche.* 2017;29(2):127-128.
19. Zartarian V, Xue J, Tornero-Velez R, Brown J. Children's Lead Exposure: A Multimedia Modeling Analysis to Guide Public Health Decision-Making. *Environ Health Perspect.* 2017;125(9):097009.
20. Linka K, Peirlinck M, Kuhl E. The reproduction number of COVID-19 and its correlation with public health interventions. *Computational Mechanics.* 2020;66(4):1035-1050.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

21. Hu Y, Yu S, Qi X, Zheng W, Wang Q, Yao H. An overview of multiple linear regression model and its application. *Zhonghua yu Fang yi xue za zhi [Chinese Journal of Preventive Medicine].* 2019;53(6):653-656.
22. Wakefield J. Multi-level modelling, the ecologic fallacy, and hybrid study designs. *Int J Epidemiol.* 2009;38(2):330-336.
23. Haldane V, Chuah FLH, Srivastava A, et al. Community participation in health services development, implementation, and evaluation: A systematic review of empowerment, health, community, and process outcomes. *PLOS ONE.* 2019;14(5):e0216112.
24. World Health Organization. Primary health care: report of the International Conference on primary health care. Paper presented at: Alma-Ata, USSR, 6-12 September 1978/jointly sponsored by the World Health Organization and the United Nations Children's Fund1978.
25. Nathan NL. *Electoral politics and Africa's Urban transition: class and ethnicity in Ghana.* Cambridge University Press; 2019.
26. Currie DJ, Smith C, Jagals P. The application of system dynamics modelling to environmental health decision-making and policy - a scoping review. *BMC Public Health.* 2018;18(1):402.
27. Factor R, Kang M. Corruption and population health outcomes: an analysis of data from 133 countries using structural equation modeling. *International journal of public health.* 2015;60(6):633-641.
28. Israëls A. Methods of standardisation. *The Hague/Heerlen, The Netherlands: Statistics Netherlands.* 2013.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Appendix

List of health care providers by category and spatial scope

Health Care Provider	Category	Spatial Scope
Clínica Norte SA	Hospital	Level 3
Clínica Medical Duarte	Hospital	Level 3
Clínica San Jose De Cúcuta SA	Hospital	Level 3
Clínica Santa Ana SA	Hospital	Level 3
Ese Hospital Universitario Erasmo Meoz	Hospital	Level 3
Uprec Unidad De Prevención Clínica - Unidad De Prevención Clínica	Intermediate	Level 3
Sanidad Policía Nacional	Intermediate	Level 3
Clínica Medico Quirúrgica Ltda	Intermediate	Level 3
Corporación Ips Norte San Rafael	Intermediate	Level 3
Centro Medico De Urgencias San Rafael Ltda	Intermediate	Level 3
Corporación Ips Norte De Santander El Parque	Intermediate	Level 3
Servicios De Salud Ips Suramericana SA	Intermediate	Level 3
UBA Vihonco - SAS	Intermediate	Level 3
UBA Vihonco Gran Colombia	Intermediate	Level 3
Comfaoriente Ips	Intermediate	Level 3
Sanaty Ips - SAS	Intermediate	Level 3
Clínica Santa Ana Sa - Sede Norte	Intermediate	Level 3
Clínica Urgencias La Merced - SAS	Intermediate	Level 3
Fundación Virgilio Barco	Intermediate	Level 3
Unidad Básica Puente Barco Leones	Intermediate	Level 2
Unidad Básica Comuneros	Intermediate	Level 2
Unidad Materno Infantil La Libertad	Intermediate	Level 2
Policlínico Atalaya	Intermediate	Level 2
Clínica Esimed La Salle	Intermediate	Level 2
Fundación Medico Preventiva Para El Bienestar Social	Primary/GP	Level 3
Clínica Medico Quirúrgica SA	Primary/GP	Level 3
Aliados En Salud SA	Primary/GP	Level 3
Norte Salud	Primary/GP	Level 3
Alvaro Ernesto Ramirez Morelli	Primary/GP	Level 3
Corporación Norte De Santander - Ips Cero	Primary/GP	Level 3
Angel Alcides Barba Barba	Primary/GP	Level 3
Ladmedis Ips Atalaya	Primary/GP	Level 3
Ramon Alberto Sánchez Carballo	Primary/GP	Level 3
Corporación Mi Ips Cúcuta Este	Primary/GP	Level 3
Salud Social Intervención Del Riesgo	Primary/GP	Level 3
Vidamedical Ips - SAS	Primary/GP	Level 3
Servicio De Asistencia Médica Inmediata SA	Primary/GP	Level 3
Corporación Ips Norte De Santander Ips Caobos	Primary/GP	Level 3
Centro Especializado De Diagnostico Materno Infantil Cedmi	Primary/GP	Level 3
Profesionales De Especialidades Medicas - SAS	Primary/GP	Level 3

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Asociación De Prestadores De Servicios Y Suministros	Primary/GP	Level 3
Diego Andres Blanco Fuentes	Primary/GP	Level 3
Hernando Antonio Villamizar Gomez	Primary/GP	Level 3
Corporación Ips Norte De Santander Ips Lleras	Primary/GP	Level 3
Eduardo Jose Fajardo Jaimes	Primary/GP	Level 3
Ana Teresa Govin Cid	Primary/GP	Level 3
Alvaro Adrian Barrera Prada	Primary/GP	Level 3
Maria Liliana Dorado Ilera	Primary/GP	Level 3
Asociación Probienestar De La Familia Colombiana -	Primary/GP	Level 3
Javier Gonzalo Corona Bueno	Primary/GP	Level 3
Cruz Roja Colombiana – CRE	Primary/GP	Level 3
Dumian Medical - SAS	Primary/GP	Level 3
Ladmedis Ips Quinta Oriental	Primary/GP	Level 2
Ips Medcare De Colombia - SAS	Primary/GP	Level 2
UBA Agua Clara	Primary/GP	Level 2
Miguel Eduardo Barrios Castillo	Primary/GP	Level 2
Central De Especialistas	Primary/GP	Level 2
Jairo Enrique Ascencio Gutiérrez	Primary/GP	Level 2
Jorge Isaac Villabona Pérez	Primary/GP	Level 2
Ips Los Olivos	Primary/GP	Level 2
Integrados En Salud Ips Unidad Básica Atalaya	Primary/GP	Level 2
Edgar Acevedo Puerto	Primary/GP	Level 2
Gerson Paris Salazar	Primary/GP	Level 2
Ips Maria De Los Angeles	Primary/GP	Level 2
Ips Orthocredit Eu	Primary/GP	Level 2
Futumedica Plus Ns - SAS	Primary/GP	Level 2
Medinorte Cúcuta Ips - SAS	Primary/GP	Level 2
Loma De Bolívar	Primary/GP	Level 1
Toledo Plata	Primary/GP	Level 1
Aeropuerto	Primary/GP	Level 1
Bocono	Primary/GP	Level 1
San Martin	Primary/GP	Level 1
Belisario	Primary/GP	Level 1
La Floresta	Primary/GP	Level 1
Buena Esperanza	Primary/GP	Level 1
San Faustino	Primary/GP	Level 1
Divina Pastora	Primary/GP	Level 1
Cundinamarca	Primary/GP	Level 1
Ospina Pérez	Primary/GP	Level 1
Belén	Primary/GP	Level 1
San Luis	Primary/GP	Level 1
Rodeo	Primary/GP	Level 1
Claret	Primary/GP	Level 1
Niña Ceci	Primary/GP	Level 1

medRxiv preprint doi: https://doi.org/10.1101/2022.04.18.22274002; this version posted April 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Contento	Primary/GP	Level 1
La Hermita	Primary/GP	Level 1
San Mateo	Primary/GP	Level 1
Guaimaral	Primary/GP	Level 1
Salado	Primary/GP	Level 1
Santa Ana	Primary/GP	Level 1
Domingo Pérez	Primary/GP	Level 1
Pórtico	Primary/GP	Level 1
Cerrito	Primary/GP	Level 1
Sevilla	Primary/GP	Level 1
Banco De Arena	Primary/GP	Level 1
Palmeras	Primary/GP	Level 1
Guaramito	Primary/GP	Level 1
Palmarito	Primary/GP	Level 1