

## Article

# Scaling dynamics of human diseases and urbanization in Colombia

Feged-Rivadeneira, Alejandro<sup>1\*</sup>; Andrade-Rivas, Federico<sup>2,3</sup>; González-Casabianca, Felipe<sup>1,5</sup>; Escobedo, Francisco J.<sup>4</sup>;

<sup>1</sup> Universidad del Rosario, Urban Management and Development, Bogotá, Colombia

<sup>2</sup> University of British Columbia, School of Population and Public Health, Vancouver, Canada

<sup>3</sup> Universidad El Bosque, Vicerrectoría de Investigaciones, Bogotá, Colombia

<sup>4</sup> USDA Forest Service, Pacific Southwest Research Station, Riverside CA, USA.

<sup>5</sup> Universidad de los Andes, Department of Mathematics, Bogotá, Colombia..

\*Corresponding Author: alejandro.feged@urosario.edu.co; Universidad del Rosario, Calle 12c No. 6-25, Bogotá, Colombia.

**Abstract:** Colombia has one of the largest numbers of internally displaced populations in the world and recently entered a period of post-conflict. These socio-political processes and trends have increased the migration of people towards cities and accordingly are affecting the distribution and occurrence of tropical diseases in its urban and peri-urban areas. Studies have suggested that many human phenomena such as urbanization scale according to the size of human populations regardless of cultural context. But other studies show that health epidemics such as malarial and human immunodeficiency virus infections, follow a scale-free distribution in terms of population size and density. We explore these relationships and dynamics in a tropical context using statistical analyses and available geospatial data to identify the scale dynamics between urbanization processes and disease transmission in Colombia. We found that rural populations and certain disease dynamics were described by power-laws that are frequently mentioned in urbanization studies. However, we found that malaria presented higher intensity of infection in human settlements of less than 50,000 individuals, particularly for ethnic indigenous populations. Results indicate that epidemics and urbanization dynamics do indeed follow scales in Colombia; findings that differ from previous epidemiological studies such as those for malarial infection. Additionally, we identified trends showing that malarial infections become endemic in peri-urban areas. Targeting such peri-urban locations and certain demographic groups are key for managing public health issues in the urbanizing tropics.

**Keywords:** Spatial Analyses; Epidemiology; Power laws; Tropical cities

## 1. Introduction

As the world is becoming increasingly urbanized, there is growing attention and uncertainty as to the potentially positive and negative effects of this process on the health of human populations. Of particular concern in middle- and low-income countries in the tropics is that urban planning and public policy are seldom informed by evidence-based population health information and there is limited monitoring to assess trends in infectious and non-communicable diseases across space and human settlements [1–3]. The potential effects of urbanization on human health might be exacerbated in the context of rapid and unplanned urbanization. One notable example is the case of Colombia, which has 75% of its population living in cities and is projected to rapidly increase in the near future [4]. Moreover, the country has one of the largest numbers of forced, internally displaced populations in the world and recently entered a period of post-conflict after more than 50 years of war between the government and The Revolutionary Armed Forces of Colombia guerrilla (FARC)

while receiving refugees and immigrants from adjacent Venezuela. In addition to other socio-ecological and geo-political factors, these population dynamics are interconnected with processes such as urbanization, deforestation, and overall land use-cover change [5]. For example, the migration of rural populations to small, medium and larger cities has resulted in the past in natural areas being converted to agricultural and urban land use covers [6].

This urbanization process and increased presence of displaced and recently arrived people in urban and peri-urban areas of Colombian cities is characterized by highly disturbed environments, poor infrastructure, and marginalized conditions [4]. As such, highlighting population density, mobility patterns, urban morphology and other socio-ecological factors in cities is key for the study of infectious disease in rapidly changing environments [7]. In Colombia, previous research conducted by some of the authors of this study showed that these factors affect public health and lead to increased vulnerability to disease burden particularly in peri-urban locations and among recently arrived ethnic minorities [8]. Even though our understanding of the connection between urbanization and population health in rapidly changing countries like Colombia is still in its early stages, it is likely that simple linear associations do not offer an adequate description of the disease burden distribution in different populations across levels of urbanization [7]. For example, while rural human settlements in early stages of the urbanization process are experiencing active disturbances such as illegal mining and deforestation, leading to higher than average risks for malaria disease occurrence; improved infrastructure in later stages of urbanization can decrease morbidity of this infection [8]. Thus, a further exploration that accounts for the complex relation between morbidity dynamics and how they change across different populations and levels of urbanization is needed.

One emergent approach is the application of scaling dynamics analyses to better understand organisms and their metabolic rates, but it can also be applied to study other phenomena, such as cities, organizations, social networks and overall sustainability [9–13]. While analyzing urban phenomena using relationships akin to those of living organisms and metabolic rates is not novel [9], such approaches have varying acceptability in the field of urban studies [10,11,13]. For example, studies on urbanization using scaling dynamics have offered insights to expand our understanding of urbanization and that several human phenomena can scale according to population size [13–16]. An emergent body of literature has also focused on the growth processes and spatial distribution of cities according to certain patterns. For example studies [14,15] have found that such patterns - regardless of cultural context and political ecologies - are bounded by factors such as population growth, energy transformation and biogeochemical cycles [17]. Therefore, understanding the scaling dynamics between urbanization and disease processes is necessary for identifying the spatial patterns of occurrence and improving public health interventions to reduce morbidity in growing urban and peri-urban areas in the tropics [6,18]. Another body of literature has also found contrasting relationships among urban population size, urbanization rates, and infectious disease occurrence such as malaria or dengue [6,19–22]. While some research suggests that urbanization leads to better access to health care and niche reductions for transmission factors, thus reducing morbidity [19–21], other studies concluded that disease occurrence might increase due to urbanization and other socio-ecological disturbance processes, as documented for disease incidence in sub-Saharan Africa [23,24].

Moreover, scaling relationships have been used to analyze several epidemic processes and their relationship to population size [25–28]. For example, measles virus infection has been reported to follow a power law in isolated populations, and available epidemiological models have failed to predict pathogen dynamics on a wider spatial scale [27]. Other studies have documented how the structure of social contacts, often described as scale-free networks, can actually determine the behavioral characteristics of an epidemic [29,30]. This is the case of sexually transmitted diseases where studies have shown that the distribution of sexual contacts scales according to population size and density [31,32]. However, the relationship between population size and disease occurrence is not yet fully understood, and results may vary depending on the health outcome of interest. While some studies have found that the relationship is inverse, meaning that as population increases the disease burden fades [19–21], other studies have concluded that population size increase and disease burden occurrence are proportional [14,15]. Previous studies in Colombia, [8,33,34] have shown that the epidemiology of malarial infection presents an abnormal distribution among indigenous groups in cities with less than 50,000 inhabitants. It is important to highlight that there is an ongoing discussion on the importance of population size on disease distribution. When looking at infectious diseases, other research has stated that population size is only one of a number of factors that determine the size of epidemics for multiple species and pathogens [35–38].

Although, recent studies on the scaling of biological phenomena continue to present consistent findings across a wide variety of species that in general organisms conform to such distributions [39,40], some argue the need for approaches to include other factors to explain disease distribution. For example, Glazier [17] suggests that the relationships between population size and diseases are insufficient given Tinbergen's [41] four laws describing biological phenomena and that such a distribution is the result of the complex interaction of diverse adaptations describing both physical, chemical, and ecological constraints.

Power laws and scale-free networks are also another approach for describing scaling dynamics in urbanization and disease related processes and phenomena [14,15]. Power laws, or scale-free distributions, are defined as a distribution of the form:  $y(x) \propto x^{-\lambda}$  for  $x > x_0$  [42], where scalar invariance produces the characteristic linear relationship between  $y$  and  $x$  variables in log-log scales. In such scale-free distributions, extreme observations are far more likely than in other types of distributions (i.e., "heavy tails") and they often lack a well-defined mean value [42]. Such scalar relationships between human and natural phenomena have been historically documented [43,44]. Studies have used such power-scaling relationships to analyze different urban phenomena such as wages, jobs, walking speed, and spreading of infectious diseases and how it can vary in relation to population size [14,15]. It is important to highlight that adequate statistical tools (e.g., maximum-likelihood analysis versus ordinary least square test) are needed to properly identify power-law distributions, as relationships between variables could be misclassified [25]. Such incorrectly fitted power-law distributions can overestimate the occurrence of large and rare observations (i.e. "heavy-tails").

The above literature supports the need to better understand the relationship between scaling, urbanization, population size, and disease processes in human populations across tropical urbanizing contexts. However, [8], there is still a gap in research on the effect of population size, and other influential factors, on pathogen loads in tropical areas with rapidly changing peri-urban land uses and socio-political processes. This is also true for chronic diseases and environmental health related disease burden.

Such information could be used to assess the relationships between the morbidity (e.g., malarial infection, chemical intoxication) of certain demographic groups, urbanization, and population size-densities in urbanizing areas of the tropics. This knowledge and application of scale-dynamic relationships has been used to aid models and decision making in various applied fields such as finance [45] and computer science [41]. Moreover, there is emergent work exploring the scaling dynamics of infection [46], and diseases such as dengue [47] and COVID-19 [48], thus showing the potential for this analytic strategy to be applied to understand other public health issues and inform policies and interventions.

In this study we explore the role of scale and its application in understanding the dynamics between urbanization and disease dynamics in Colombia. To better understand such scale dynamics among these processes we test whether population size and epidemic characteristics have similar scales that follow a conventional power-law distribution. In addition, to explore the nuances of scaling dynamics for particular diseases, we used malaria as a case study to further assess the relationship between population growth and disease occurrence. We test for these objectives with an integrated approach using likelihood ratios and hypotheses tests. Assessing the role of these changing population sizes in terms of disease outcomes and demographics is key for public health surveillance and preventative planning in urbanizing tropical contexts such as Colombia.

## 2. Materials and Methods

### 2.1 Study area

Colombia has a population of about 50 million people and is one of the most biodiverse countries in the world with a wide range of socio-ecological contexts due to a diversity of biomes, elevations, temperature and precipitation regimes and other socio-ecological factors [5]. We included all 1,222 municipalities (second level administrative division after departments). Despite land consumption being relatively low in Colombia, urban density in some cities are among the highest in the world, and there is a tendency to urbanize environmentally hazardous areas [4,49]. The dynamics related to urbanization and other land cover changes such as deforestation are complex and include a number of factors, such as national and local level governance and governability issues, socio-political instability, climate change effects, inequities, and on-going low level conflict from armed groups [5,49]. This long history of social conflict, inequitable distribution

of wealth, lack of land tenure, poverty, and lack of governability in the country have led to decades of armed conflict and subsequent socio-political unrest. Nevertheless, with the 2016 signing of a peace accord with the FARC guerrilla, there is speculation as to what will happen in terms of LUC such as deforestation, urbanization and population-demographic shifts.

## 2.2 Data

We used case report data from the National Health Institute (INAS in Spanish) for all 74 diseases reported in the national surveillance system during 2007- 2015 as well as available population data for 1,122 different municipalities from the National Department of Statistics (DANE, acronym in Spanish). All 2,455,617 compiled case reports specify type of disease, occurrence locations at the municipality level, and demographic variables such as age and ethnicity. Individual-level case reports were not available; however, aggregated anonymous data ( $n = 2,455,617$ ) were obtained from a cube-query system maintained by the Social Health Protection Ministry (Ministerio de Protección Social, 2010). We aggregated cases at the municipality scale according to ethnicity and age, since the intensity of the disease has been reported to vary across different ethnicities [33,34]. The three following ethnic groups were analyzed: Afro-Colombian (AFRO), Indigenous (IND), and No-ethnic designation (ND). We focused on these populations as they represent over 99% of the total case reports and we categorized disease events in either chronic transmission, vector borne or environmental subgroups. We note that some diseases are associated with more than 1 event such as malaria and complicated malaria.

In the subsequent analyses we focused on the following events, grouped into four different types (Table I) :

Type	Events
Environmental	snakebite, exposition to flour, adverse event following vaccination, aggression by animal potentially infected by rabies, acute respiratory infection, rare acute respiratory infection, intoxication with plague control chemicals, intoxication with drugs (medicines), intoxication with methanol, intoxication with heavy metals, intoxication with solvents, intoxication with other substances, intoxication with carbon monoxide and other gases, intoxication with psychoactive substances, injury by explosive artifact, injury by fireworks, congenital anomalies, domestic violence
Non Communicable	low birth-weight, death due to malnutrition, congenital malformations, congenital hypothyroidism, acute lymphoid pediatric leukemia, acute myeloid pediatric leukemia, typhoid and paratyphoid fever, infant cancer, maternal mortality, perinatal and neonatal mortality, mortality due to acute diarrhea 0-4 years old, mortality due to acute respiratory infection,
Communicable	hepatitis A, B and C, Leprosy, leptospirosis, meningitis (meningococcica), meningitis (Haemophilus influenzae), meningitis (pneumococcus), meningitis (tuberculosis), meningitis (bacterial), parotitis, measles, acute diarrhea due to rotavirus, congenital syphilis, gestational syphilis, pertussis, tuberculosis (extra-pulmonary), tuberculosis (pulmonary), tuberculosis (resistant), chickenpox, HIV
Vector Borne	chagas disease, dengue, complicated dengue, chikungunya, cutaneous Leishmaniasis, mucosal Leishmaniasis, visceral Leishmaniasis, malaria (mixed), malaria (P.falciparum), malaria (P.vivax), malaria (P.malariae), malaria (complicated), mortality due to malaria, mortality due to dengue, epidemic typhus transmitted by lice, endemic typhus transmitted by fleas

Table I Events considered with their corresponding type.

## 2.3 Population, epidemics and power-law distribution scales analysis

We tested our objective using the Kolmogorov-Smirnoff (KS) statistic to calculate the difference among observed, theoretical, and truncated distributions. A truncated distribution finds the value of  $X_{min}$  such that the observed distribution minimizes the value of the KS statistic [50]. The KS test also simulated numerous data sets and compared each occurrence with the observed distribution. We used a significance level of 0.10, performed the test on each simulated distribution, and counted the number of instances when the null hypotheses were rejected ( $H_0$ : two data sets came from the same distribution). We determined the power-law to be a good fit for the observed data if the null hypothesis was rejected in less than 10% of the simulations. To examine if the other distributions

also fit the data, we implemented the same test for an exponential and log-normal distribution. We performed KS tests for three distributions (i.e., power-law, exponential, log-normal) for both population and disease incidence separately, and we subsequently combined them. The KS test was also performed independently for each of the three analyzed ethnic groups. All statistical tests and spatial mapping procedures were analyzed using the R statistical software package [51], specifically the distribution test code in [52].

### 3. Results

A total of 2,455,617 case reports were analyzed for the period 2007-2015. Table II shows the distribution of total case reports by year and ethnicity. Figure 1 illustrates the time trend of four disease types (i.e, non-communicable, environmental, vector borne, communicable) showing that some disease events increase over time, probably following the country's population growth trends (population time series data were not available for a further assessment). Within each category, chickenpox, newborn mortality, and dengue presented a relatively high number of cases during the study period, while for environmental events several diseases had relatively high counts. Figure 2 shows the spatial distribution of diseases by category, where chronic disease events per thousand people are higher among the most densely populated areas of the country, and diseases classified as transmission events present a higher rate along the Pacific Coast, the Caribbean Coast, and the Amazon Basin.

Year	Afro	Indig.	N.D	Palenquero	Raizal	Rom	Total
2007	12991	2999	66627	17	210	161	83005
2008	19438	4924	118113	44	1336	492	144347
2009	20418	6297	187411	114	1379	594	216213
2010	31342	8948	276671	149	2421	662	320193
2011	23050	6490	181823	181	670	320	212534
2012	24229	10783	268433	276	728	358	304807
2013	27892	12496	314429	262	592	217	355888
2014	31145	13065	343884	189	694	1918	390895
2015	38038	18600	368483	132	707	1775	427735
Total	228543	84602	2125874	1364	8737	6497	2455617

Table II Case reports in Colombia for 2007-2015 by ethnicity and year. Note that the total cases for the Indigenous, Afro-Colombian and Other (N.D.) comprise over 99% of the total case reports.



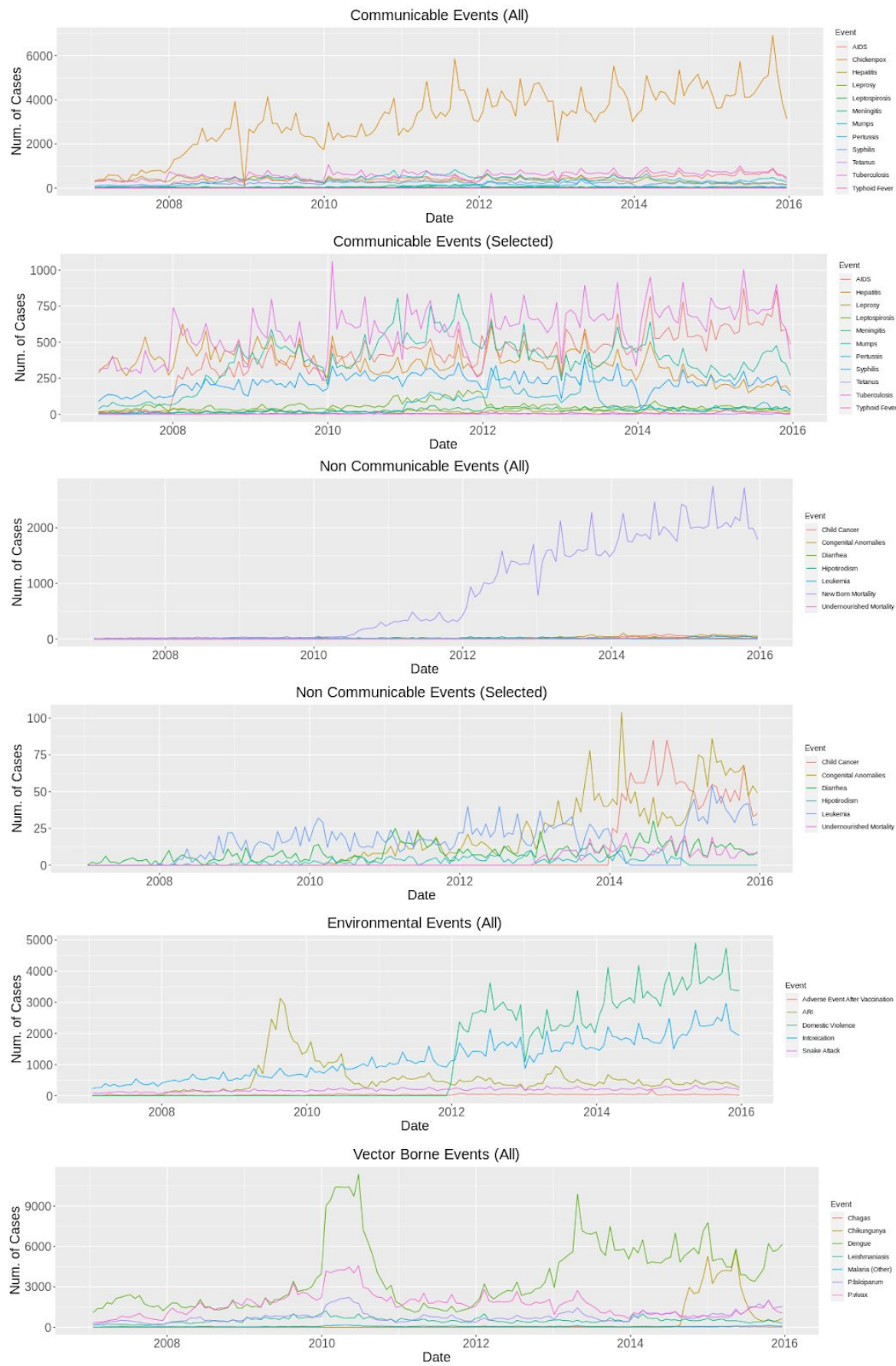


Figure 1 Case reports events by week and disease type over the study period in Colombia.

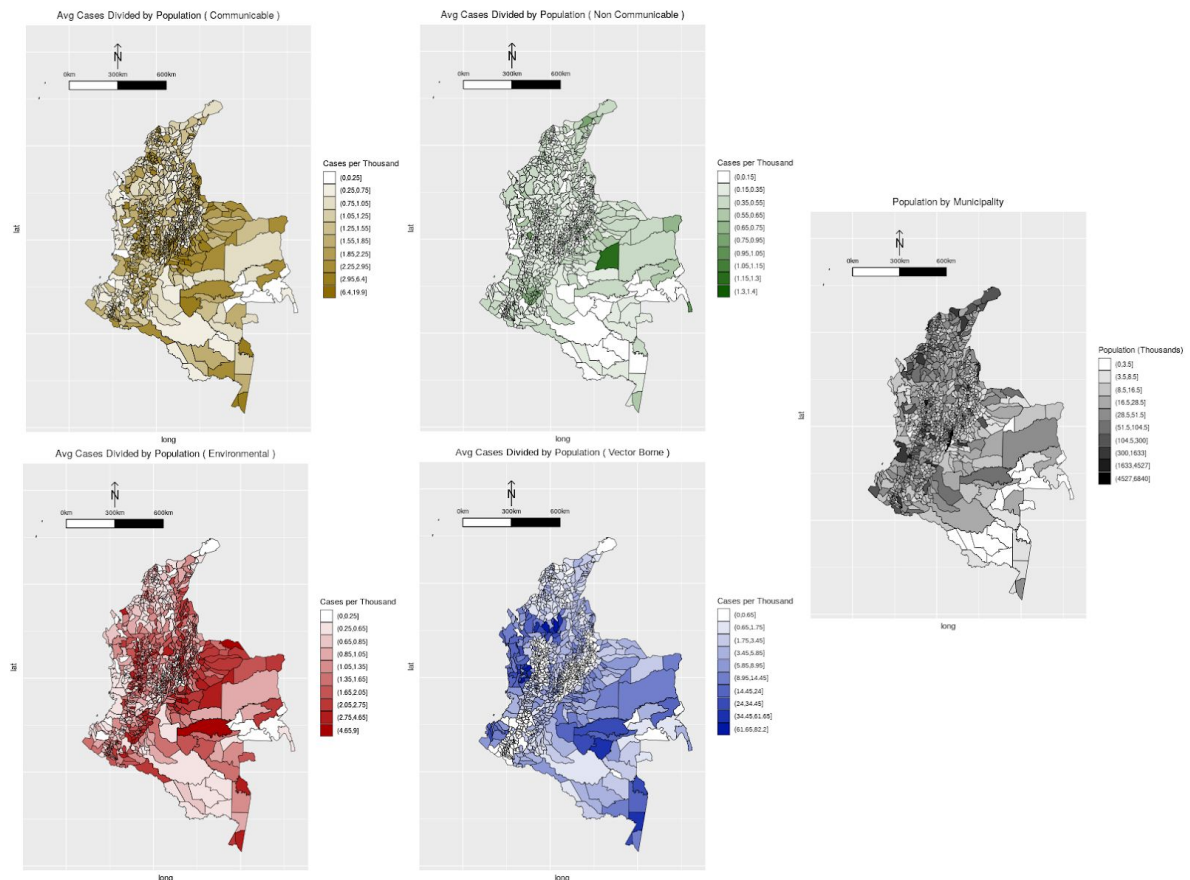


Figure 2 Average case reports events by municipality in Colombia. Events have been aggregated into categories for the purpose of visualization (i.e. communicable, non communicable, environmental and vector borne).

### 3.1 Population size, epidemics and power-law scale distributions

The complete table of results is included in the supplementary materials.

Based on the KS test for power-law distribution, we found that the total population ( $x_{\min} = 14784$ ;  $\alpha = 2.15$ , KS-test rejection= 2.80%), urban population ( $x_{\min} = 8425$ ;  $\alpha = 1.86$ , KS-test rejection= 0.80%) and rural ( $x_{\min} = 18252$ ;  $\alpha = 3.53$ , KS-test rejection= 1.52%) scaled according to the power-law distribution. However, our test also showed that both all and urban populations follow an exponential distribution and rural a log-normal distribution.

When applying the KS test for the prevalence (Case reports events/Population) of all the diseases included in this study, we found it scales according to a power-law ( $x_{\min} = 0.05$ ;  $\alpha = 2.78$ ; KS-test rejection= 1.68%). In addition, the test results according to ethnic groups showed that total disease prevalence and ethnic population followed a power-law for the ND ( $x_{\min} = 0.06$ ;  $\alpha = 3.2$ ; KS-test rejection= 2.04%), and AFRO group ( $x_{\min} = 0.00$ ;  $\alpha = 1.64$ ; KS-test rejection= 2.56%). However, it was not the case for the IND group ( $x_{\min} = 0.00$ ,  $\alpha = 1.45$ ; KS-test rejection= 25.8%). Interestingly, this group did show a power-law distribution when focused only on the rural population ( $x_{\min} = 0.01$ ,  $\alpha = 2.00$ ; KS-test rejection= 1.04%). The calculated alpha values for the ND group was the highest, followed by the IND class, and then the AFRO group which had the lowest alpha. This suggests that as populations scales, the ND population has the largest reduction in overall disease prevalence, followed by the IND population (in the rural population), and the AFRO with the least effect associated with population growth.

A further analysis of scaling dynamics was also conducted for the four established disease categories (i.e., communicable, non-communicable, environmental, and vector borne). Figure 3 illustrates the power-law alpha values compared to the prevalence of the four disease types and

malaria, by ethnic class. The KS test showed that communicable diseases for the ND ethnic class scaled according to the power-law distribution ( $x_{\min} = 0.02$ ;  $\alpha = 4.40$ ; KS-test rejection= 1.32%), and did not scale according to exponential and log-normal distributions (both KS-test rejection = 0.0%). In contrast, communicable diseases did not scale for IND and AFRO classes. For the case of non-communicable diseases we found that the total population ( $x_{\min} = 0.00$ ;  $\alpha = 5.51$ ; KS-test rejection = 2.32%), ND ( $x_{\min} = 0.00$ ;  $\alpha = 5.39$ ; KS-test rejection = 3.80%), and AFRO ( $x_{\min} = 0.00$ ;  $\alpha = 3.4$ ; KS-test rejection = 0.90%) scaled according to a power-law distribution, while the test showed that this was not the case for log-normal and exponential distribution. The IND ethnic class showed a similar trend ( $x_{\min} = 0.02$ ;  $\alpha = 3.2$ ; KS-test rejection = 1.00%). In line with the results for the total of diseases analysed, the ND category presented the highest alpha value, showing that this population experiences a greater reduction of non-communicable diseases as population grows.

The KS test results for environmental diseases show evidence that the prevalence of environmental diseases for the total population scaled and followed a power-law distribution ( $x_{\min} = 0.02$ ;  $\alpha = 5.54$ ; KS-test rejection= 0.70%), and did not follow a log-normal or exponential distribution (KS-test rejection close to zero percent). We found that environmental diseases among ND populations scaled according to a power-law ( $x_{\min} = 0.01$ ;  $\alpha = 4.43$ ; KS-test rejection= 2.52%) and did not follow either a log-normal or an exponential distribution. In contrast, there is no evidence that this is the case for AFRO ( $x_{\min} = 0.00$ ;  $\alpha = 1.68$ ; KS-test rejection= 75.7%) and IND populations ( $x_{\min} = 0.00$ ;  $\alpha = 1.60$ ; KS-test rejection= 36.9%). In the last disease category (i.e. vector borne diseases), the KS test showed clear evidence that the total population ( $x_{\min} = 0.03$ ;  $\alpha = 2.25$ ; KS-test rejection= 2.4%), as well as the IND ( $x_{\min} = 0.00$ ;  $\alpha = 1.46$ ; KS-test rejection= 2.88%), AFRO ( $x_{\min} = 0.00$ ;  $\alpha = 1.54$ ; KS-test rejection= 2.08%), and ND ( $x_{\min} = 0.04$ ;  $\alpha = 2.75$ ; KS-test rejection=1.96%) classes, scaled according to a power-law distribution. Moreover, the tests for log-normal and exponential distributions, for the total population and across ethnic classes, were far below our KS test rejection percentage. Consistent with these findings for the other disease groups, the ND class presented the highest alpha among ethnic classes, thus providing evidence to support that larger human settlements are associated with more reduction in vector borne diseases for ND populations when compared to the AFRO and IND ethnic classes.



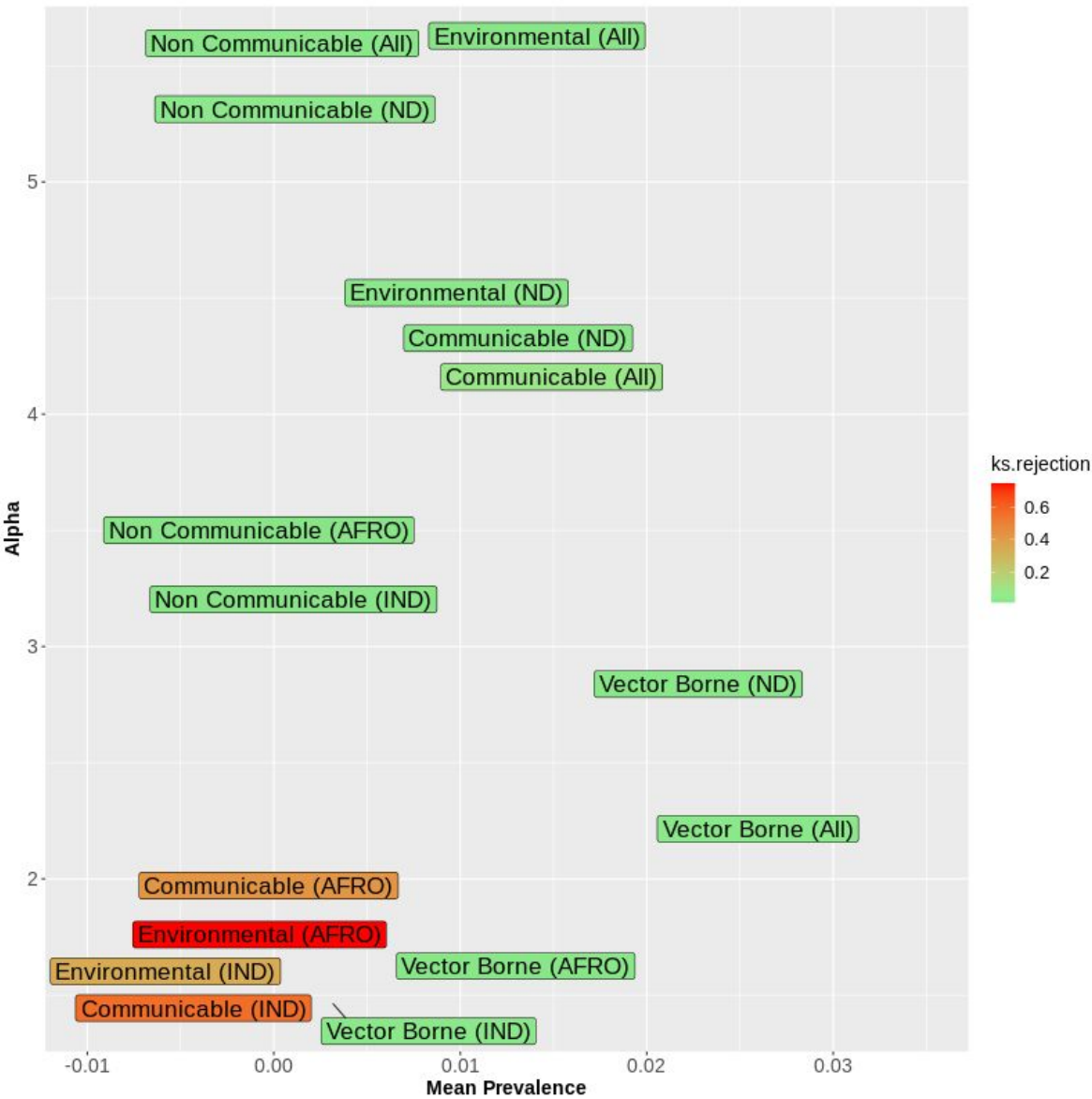


Figure 3 Alpha values for the power-law tests and incidence values for different diseases and public health issues in Colombia, grouped by type and ethnicity.

In general, all of the diseases included in this study scaled as a power-law, with leukemia and domestic violence having the highest rejection values (7%). Figure 4 shows the alpha values for the power-law tests compared to the disease prevalence, for all diseases and public health issues considered. It is worth noticing the effect of population size varies depending on the disease. For example, malaria is the least affected from increases in population size. A limited effect of population size on the disease events was also found for Chagas, diarrhea, leptospirosis and leishmaniasis. In contrast, newborn mortality, dengue, chickenpox, intoxication and AIDS, presented the highest alpha values, showing a relatively high effect of population size on disease prevalence when compared to other event cases (Figure 4).

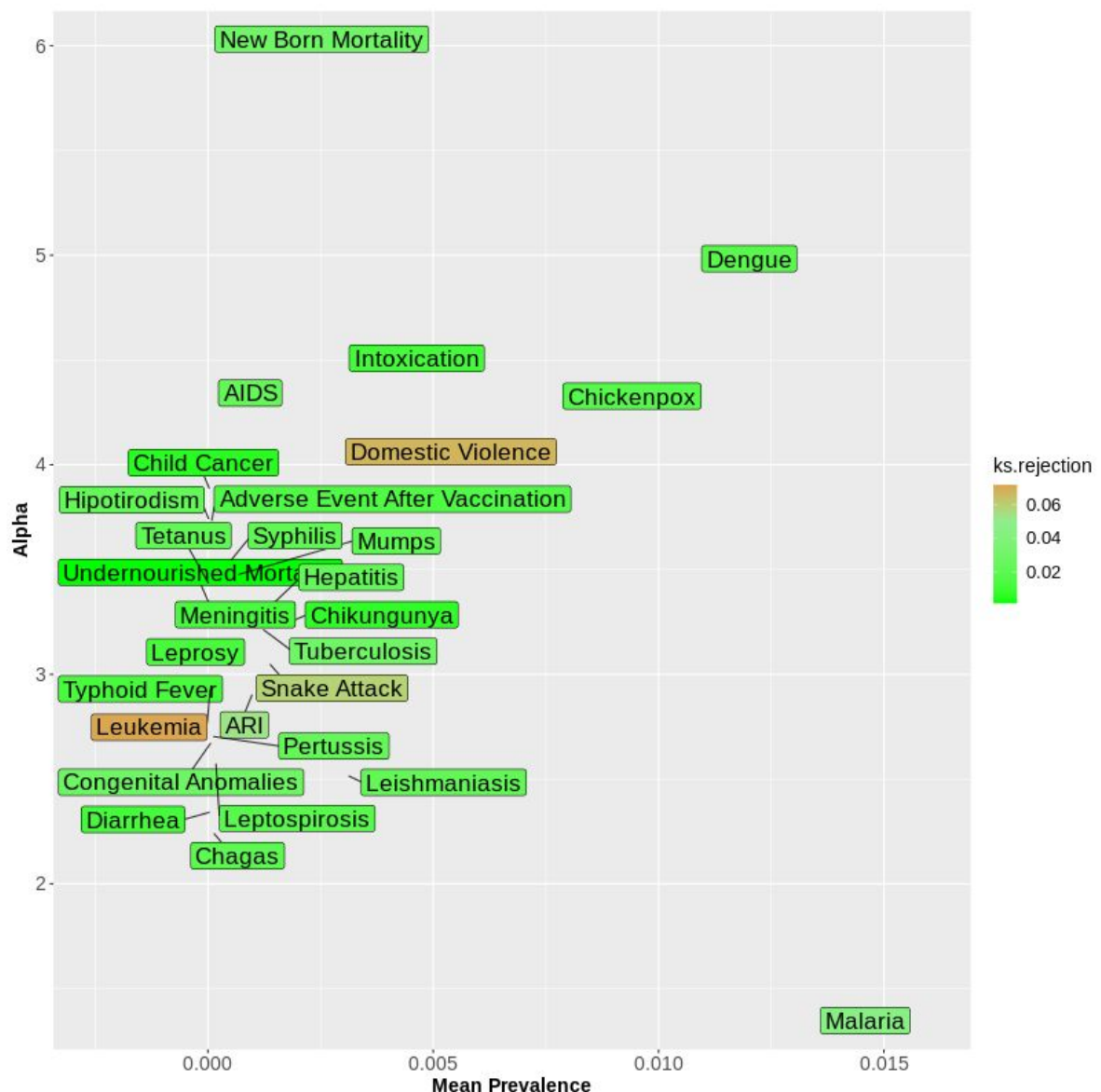


Figure 4 Alpha values for the power-law tests and incidence values for different diseases and public health issues in Colombia. Note: ARI, Acute respiratory infections.

In order to further explore how disease occurrence may be associated with population growth and urbanization dynamics, an assessment of population size, and ethnicity was conducted for malaria infection. In recent years, the increasing trend in malaria cases has raised concern among researchers and public health authorities and is jeopardizing the goals of the National Malaria Control Programme of eliminating urban malaria by 2021 [8,53,54]. Hence, we conducted a series of KS tests for malaria prevalence and found evidence of disease events scaling and following a power-law distribution for the total population ( $x_{\min} = 0.00$ ;  $\alpha = 1.43$ ; KS-test rejection= 4.48%), the AFRO ethnic class ( $x_{\min} = 0.00$ ;  $\alpha = 1.36$ ; KS-test rejection= 3.88%) and the ND ( $x_{\min} = 0.00$ ;  $\alpha = 1.54$ ; KS-test rejection= 3.48%). Moreover, we found no evidence that malaria prevalence follows a log-normal or exponential distribution for the total population and each of the ethnic categories. A relevant disease occurrence metric to understand malaria distribution is the intensity of disease (cases less than 5 years-old/total cases). When conducting the KS test for malaria intensity we found a greater effect of population on events occurring for the total population ( $x_{\min} = 0.02$ ;  $\alpha = 2.41$ ; KS-test rejection= 1.40%) when compared to our results looking at the scaling dynamics for malaria prevalence. The ND ( $x_{\min} = 0.00$ ;  $\alpha = 1.80$ ; KS-test rejection= 1.84%), and IND ( $x_{\min} = 0.12$ ;  $\alpha = 3.37$ ; KS-test rejection= 3.20%) ethnic classes also showed evidence that malaria intensity scales following a power-law distribution, while the AFRO class was just below the 5% rejection percentage ( $x_{\min} = 0.01$ ;  $\alpha = 2.32$ ; KS-test rejection= 6.20%). There is no statistical

evidence that malaria intensity scales following a log-normal or exponential distribution for the ethnic classes analyzed.

Overall our results also indicated an association between malarial intensity and population size (Figure 5). We found specific ranges in population size where malaria infections actually increased. Figure 5 shows that there is an interval in population size (0-100,000) where indigenous populations experience a statistically significant greater malaria intensity when compared to ND and AFRO ethnic classes. The IND group also experienced the most acute increase in malarial infection intensity among all ethnic groups.

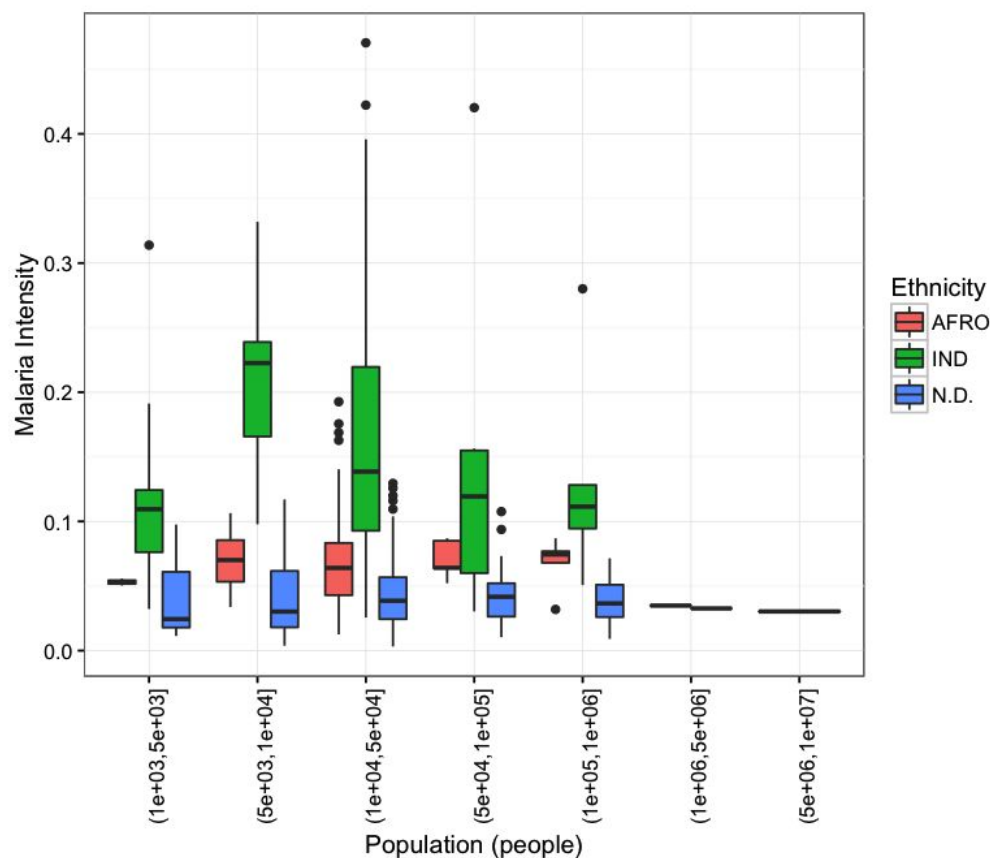


Figure 5 Malaria intensity (cases < 5 years-old/total cases) and population size, divided by ethnic groups. The graph shows that there is an interval in population size (5,000-100,000) when indigenous populations experience a significant higher malaria intensity than the other ethnic groups.

#### 4. Discussion

This study explored the scaling dynamics of population growth and disease via power law distributions. In addition, we used malaria as a case study to further assess the relationship between population growth and disease occurrence for a specific disease. Case reports during the study period presented different trends across the four disease categories used in this study (Figure 1). Similarly, the spatial distribution of case report incidences for these categories sheds some light on specific areas of public health concern, and how they varied according to different types of disease. For example, a noticeable spatial pattern of elevated cases of vector borne disease occurrence can be identified with three major geographic areas: the Pacific Coast and Antioquia (west), Amazon region (south), and the Eastern Plains (east) (Figure 2). Thus, using spatio-temporal scan statistics, a comparable distribution was confirmed by other studies in the case of malaria outbreaks in Colombia between 2007 and 2015 [8]. We note that the Easter Plains region presents consistently high rates across the four disease categories (Figure 2).

Overall, in Colombia we found that the prevalence of case reports scales following a power-law distribution among the total population. When looking at different types of diseases studied across ethnic classes, the ND ethnic class consistently scaled following a power-law distribution. Moreover, the ND group presented the highest alpha among ethnic classes for total, non-communicable, and vector borne diseases. Therefore, the effects of population growth on the reduction of disease prevalence may be stronger for populations that do not identify as Indigenous or Afrocolombian. In fact, for the IND and AFRO ethnic classes, environmental and communicable diseases did not scale with population size, suggesting that the effects of larger urban centers on disease prevalence reduction, may not be present for these particular populations in Colombia. One potential explanation is that these populations tend to live, or settle following internal migration, in urban and peri-urban areas where access to infrastructure and health care is limited, and environmental and communicable diseases drivers are not controlled or may be exacerbated [33,34]. Other social, cultural, and economic factors could impact the accessibility of these populations to safe housing and infrastructure conditions [33,34]. Further studies should be conducted to assess the distribution of disease across ethnic classes in smaller administrative units including an assessment of differences between urban, peri-urban, and rural settlements.

For the other disease types (i.e., vector borne, non-communicable), the AFRO and IND populations did show evidence of a reduction in disease prevalence, however we found a more moderate decrease than for the ND group. This sheds light on potential health equity issues where the benefits of health care facilities and other services provided in larger settlements do not have the same positive effect on IND and AFRO populations as it has among the ND group. This finding should be further explored in order to determine causal pathways, although some preliminary studies have highlighted a number of barriers these populations must face to access health care and how these populations are segregated to neighbourhoods with precarious sanitation and infrastructure [55–58]. In regards to non-communicable disease risk factors, there is evidence that income inequality and socioeconomic status in Colombia are associated with hypertension and cardiovascular and metabolic risk [59,60]. Although there is some evidence that cardiovascular and metabolic risk is not associated with ethnicity after accounting for socioeconomic status, overall ethnic differences in non-communicable diseases still remain understudied in Colombia [59].

For most of the individual health outcomes analyzed, disease prevalence and population size in Colombia do scale according to power-law properties (Figures 4). This decrease in disease burden according to population size has also been observed in previous studies by Keiser et al. [24] and Feged-Rivadeneira et al. [34]. Indeed the effect of population on disease prevalence varies across diseases and public health issues, as well as disease metrics. For example, population size seems to have a limited effect over diseases like dengue, chikungunya, respiratory infections, typhoid, and whooping cough, while tetanus, leishmaniasis, and maternal mortality showed a more pronounced decrease as population size increases (Figure 4). Moreover, some disease burden metrics, that are relevant for specific health outcomes, provide complementary information about the dynamics of disease. In our study, we found that the malaria intensity (cases less than 5 years-old/total cases) provided additional evidence of the scaling dynamics of the disease, showing a stronger association between population size and disease for the total population and across ethnic groups than for disease prevalence. Thus, in order to assess the scaling dynamics of a particular disease, other relevant metrics of disease occurrence that expand our understanding of the disease burden should be explored.

The association between population size and disease might require a more nuanced analysis, as the effect of population on disease may be mediated by other factors at particular population sizes. For example, our results also indicated an association between malaria intensity and population size (Figure 5). Previous ethnographic work [33,34] suggested a consistent pattern across Colombia where urban areas population growth is associated with a reduction in malaria infection, a pattern that was later assessed using spatiotemporal surveillance tools [8]. However, this study showed specific ranges in population size where malaria infections actually increased for Indigenous populations (Figure 5). Specifically, we found that in the interval between 5000 and 50,000 inhabitants, Indigenous populations presented a greater disease burden than expected by their population size (Figure 5). This may be related with Indigenous populations residing in peri-urban areas - where an intricate network of interactions occur. Although Indigenous populations might present some reduction in disease occurrence when living in human settlements



with larger populations, it is crucial to assess the absence of this reduction within this population size range for malaria infection. Thus establishing a robust health surveillance system to monitor disease, population size, and vulnerability in these populations is warranted in order to prevent a niche of malarial transmission infection by the parasite becoming endemic. Moreover, a further understanding of the distribution of disease in peri-urban areas across ethnic groups should be explored in future research, as some research has already shown that social-ecological factors in peri-urban areas may exacerbate health issues such as malaria and parasite infections in children among Indigenous populations [33,61,62].

Several potential limitations should be considered when interpreting our results. First, we are exploring aggregated data of a diverse nature in very complex socio-ecological systems across multiple socio-political scales in communities that are often resource poor. Second, some diseases in Colombia are more frequently reported than others, thus observed differences may be the result of reporting or data quality and increased surveillance; thus, observed differences should be interpreted with caution. That said, our integrated statistical-spatial approach to analyzing our data in an aggregated manner, should be robust to underreporting. An additional limitation is that observed trends could change if disease surveillance and reporting were to be improved. This is particularly important in the case of highly vulnerable populations since smaller populations might be underestimated if a few number of cases were not to be reported.

Our findings show that in post-conflict, tropical Colombia, factors associated with urbanization such as population size as well as ethnicity and socio-ecological contexts can create the necessary conditions and niches for disease transmission as has been reported in other studies [35,36]. Additionally, other studies have also documented how epidemic characteristics of pandemics such as communicable and vector-borne diseases change due to modifications of host populations [14,63], the approach and findings presented in this study can be used as a means to understand and create awareness of the importance of specific municipalities receiving large populations in peri-urban areas, to be cognizant and prepared for the possible impacts of urbanization-related process on disease occurrence in contexts with highly dynamic occurring natural and anthropogenic disturbances [4,37,38].

## 5. Conclusions

Findings from this study have potential implications for public health and socio-ecological disturbance research in countries such as Colombia. Researchers can use publicly available geospatial and disease occurrence data to quantify relationships that have been documented elsewhere in the world. Such data-driven quantitative analyses have some advantages over other more complex and expensive studies depending on site-specific and multiple disease transmission data based on extensive clinical and field work (e.g., trapping mosquito populations and estimating disease incidence in each location). The relationship identified in this study between population size and disease occurrence can be used to more readily map, understand and predict public health issues, assess the impact of land use and cover change, and explore the distribution of health outcomes in urban and peri-urban settings. This type of research is crucial in a country where the migration of populations to cities due to internal displacement and current environmental change (e.g., climate change) might exacerbate population health issues. Our study's findings also have public health intervention, planning and surveillance applications. First, we have documented that peri-urban locations of both small and large urban settlements have higher disease intensity, while at the same time, large urban centers experience lower burden of disease. Our findings suggest that public health intervention units should vary according to their epidemiologic characteristics in terms of ethnic groups, population size, and disease as these have specific scaling and epidemic properties. Such findings could be used to design more cost-effective public health and surveillance systems.

Research on tropical diseases like malaria infection and its relationship to urbanization in places such as Colombia can inform and improve public health interventions at the local, regional and national levels. Moreover, analysing the malaria case study shed light on potential variations on the relationship between population size and disease burden, which depend both on the type of disease and metrics used. Colombia, with its ongoing post-conflict process, accelerated

deforestation rates, urbanization processes, and history of displaced peoples and their movement to cities, presents a novel context for better understanding these dynamics. Given the high degree of variation in terms of the intensity of disease infections and the heterogenous environments created by tropical disturbance processes in low-middle income countries, Colombia offers valuable evidence to explore the dynamics of urbanization, land use, post-conflict, and disease relationships according to population size and scales.

### Supplementary Materials

The table: Complete Experiments Results, in the supplementary materials, contains all the results of the different distribution tests executed for each ethnic group and group of events.

**Acknowledgments:** This work was funded by the Colombian Directorate for the Promotion of Research, National Program in Science, Technology and Innovation in Environment, Biodiversity and Habitat (Grant 122274558511). FA-R would like to acknowledge the Canadian Institutes of Health Research for his Vanier Canada Graduate Scholarship and Colciencias (Colombia) for his International PhD Scholarship.

**Author Contributions:** “All authors have read and agree to the published version of the manuscript. Conceptualization, AFR., FE, FA, FGC; methodology, AFR, FGC.; software, FGC; validation, FGC, FA, FE; formal analysis, FGC.; investigation, AFR., FE, FA, FGC; resources, AFR.; data curation, AFR, FGC.; writing—original draft preparation, AFR, FGC.; writing—review and editing, FA, FE.; visualization, FGC.; supervision, AFR.; project administration, AFR.;”.

**Funding:** This work was funded by the Colombian Directorate for the Promotion of Research, National Program in Science, Technology and Innovation in Environment, Biodiversity and Habitat (Grant 122274558511).

**Conflicts of Interest:** The authors declare no conflict of interest..

**Data Availability:** The data that support the findings of this study are openly available in <https://figshare.com> at <https://doi.org/10.6084/m9.figshare.13146347>.

### References

1. Daniels, M.E.; Donilon, T.E.; Bollyky, T.J. The Emerging Global Health Crisis: Noncommunicable Diseases in Low- and Middle-Income Countries. *SSRN Electron. J.* **2014**, doi:10.2139/ssrn.2685111.
2. Eckert, S.; Kohler, S. Urbanization and health in developing countries: a systematic review. *World Health Popul.* **2014**, *15*, 7–20, doi:10.12927/whp.2014.23722.
3. Neiderud, C.-J. How urbanization affects the epidemiology of emerging infectious diseases. *Infect. Ecol. Epidemiol.* **2015**, *5*, 27060, doi:10.3402/iee.v5.27060.
4. Green, T.L.; Escobedo, F.J.; Arango, J.M. ¿Ciudades sostenibles o ciudades resilientes? *En Diálogo Con Tierra Por Una Colomb. Sosten.* **2017**, 75–85, doi:10.2307/j.ctv11wjh4.9.
5. Sánchez-Cuervo, A.M.; Aide, T.M. Consequences of the Armed Conflict, Forced Human Displacement, and Land Abandonment on Forest Cover Change in Colombia: A Multi-scaled Analysis. *Ecosystems* **2013**, *16*, 1052–1070, doi:10.1007/s10021-013-9667-y.
6. Wen, T.-H.; Tsai, C.-T.; Chin, W.-C.-B. Evaluating the role of disease importation in the spatiotemporal transmission of indigenous dengue outbreak. *Appl. Geogr.* **2016**, *76*, 137–146, doi:10.1016/j.apgeog.2016.09.020.
7. Santos-Vega, M.; Martinez, P.P.; Pascual, M. Climate forcing and infectious disease transmission in urban landscapes: integrating demographic and

- socioeconomic heterogeneity. *Ann. N. Y. Acad. Sci.* **2016**, *1382*, 44–55, doi:10.1111/nyas.13229.
8. Feged-Rivadeneira, A.; Ángel, A.; González-Casabianca, F.; Rivera, C. Malaria intensity in Colombia by regions and populations. *PLOS ONE* **2018**, *13*, e0203673, doi:10.1371/journal.pone.0203673.
  9. Gandy, M. Rethinking urban metabolism: water, space and the modern city. *City* **2004**, *8*, 363–379.
  10. Lefebvre, H. *The production of space*; Oxford Blackwell, 1991; Vol. 142;.
  11. Lefebvre, H. *The urban revolution*; U of Minnesota Press, 2003;
  12. West, G.B. *Scale: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies*; Penguin Press: New York, 2017; ISBN 978-1-59420-558-3.
  13. Brenner, N.; Schmid, C. Towards a new epistemology of the urban? *City* **2015**, *19*, 151–182.
  14. Bettencourt, L.M.; Lobo, J.; Helbing, D.; Kühnert, C.; West, G.B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci.* **2007**, *104*, 7301–7306.
  15. Bettencourt, L.M. The origins of scaling in cities. *science* **2013**, *340*, 1438–1441.
  16. Noulas, A.; Scellato, S.; Lambiotte, R.; Pontil, M.; Mascolo, C. A tale of many cities: universal patterns in human urban mobility. *PloS One* **2012**, *7*, e37027.
  17. Glazier, D.S. Beyond the ‘3/4-power law’: variation in the intra-and interspecific scaling of metabolic rate in animals. *Biol. Rev.* **2005**, *80*, 611–662.
  18. Pou, S.A.; Tumas, N.; Sánchez Soria, D.; Ortiz, P.; Díaz, M. del P. Large-scale societal factors and noncommunicable diseases: Urbanization, poverty and mortality spatial patterns in Argentina. *Appl. Geogr.* **2017**, *86*, 32–40, doi:10.1016/j.apgeog.2017.06.022.
  19. Hay, S.I.; Snow, R.W. The Malaria Atlas Project: Developing Global Maps of Malaria Risk. *PLOS Med.* **2006**, *3*, e473, doi:10.1371/journal.pmed.0030473.
  20. Hay, S.I.; Guerra, C.A.; Tatem, A.J.; Atkinson, P.M.; Snow, R.W. Urbanization, malaria transmission and disease burden in Africa. *Nat. Rev. Microbiol.* **2005**, *3*, 81–90, doi:10.1038/nrmicro1069.
  21. Snow, R.W.; Guerra, C.A.; Noor, A.M.; Myint, H.Y.; Hay, S.I. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **2005**, *434*, 214–217.
  22. Ng, S.C.; Shi, H.Y.; Hamidi, N.; Underwood, F.E.; Tang, W.; Benchimol, E.I.; Panaccione, R.; Ghosh, S.; Wu, J.C.Y.; Chan, F.K.L.; et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *The Lancet* **2017**, *390*, 2769–2778, doi:10.1016/S0140-6736(17)32448-0.
  23. Donnelly, M.J.; McCall, P.; Lengeler, C.; Bates, I.; D’Alessandro, U.; Barnish, G.; Konradsen, F.; Klinkenberg, E.; Townson, H.; Trape, J.-F.; et al. Malaria and urbanization in sub-Saharan Africa. *Malar. J.* **2005**, *4*, 1.
  24. Keiser, J.; Utzinger, J.; De Castro, M.C.; Smith, T.A.; Tanner, M.; Singer, B.H. Urbanization in sub-saharan Africa and implication for malaria control. *Am. J. Trop. Med. Hyg.* **2004**, *71*, 118–127.
  25. Goldstein, M.L.; Morris, S.A.; Yen, G.G. Problems with fitting to the power-law distribution. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **2004**, *41*, 255–258.
  26. Hochachka, W.M.; Dhondt, A.A. Density-dependent decline of host abundance resulting from a new infectious disease. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5303–5306, doi:10.1073/pnas.080551197.

27. Newman, M.E. Spread of epidemic disease on networks. *Phys. Rev. E* **2002**, *66*, 016128.
28. Rhodes, C.J.; Anderson, R.M. Power laws governing epidemics in isolated populations. *Nature* **1996**, *381*, 600–602, doi:10.1038/381600a0.
29. Eubank, S.; Guclu, H.; Kumar, V.A.; Marathe, M.V.; Srinivasan, A.; Toroczkai, Z.; Wang, N. Modelling disease outbreaks in realistic urban social networks. *Nature* **2004**, *429*, 180–184.
30. Salathé, M.; Vu, D.Q.; Khandelwal, S.; Hunter, D.R. The dynamics of health behavior sentiments on a large online social network. *EPJ Data Sci.* **2013**, *2*, 1–12, doi:10.1140/epjds16.
31. Beyrer, C.; Baral, S.D.; van Griensven, F.; Goodreau, S.M.; Chariyalertsak, S.; Wirtz, A.L.; Brookmeyer, R. Global epidemiology of HIV infection in men who have sex with men. *The Lancet* **2012**, *380*, 367–377, doi:10.1016/S0140-6736(12)60821-6.
32. Leventhal, G.E.; Günthard, H.F.; Bonhoeffer, S.; Stadler, T. Using an epidemiological model for phylogenetic inference reveals density dependence in HIV transmission. *Mol. Biol. Evol.* **2014**, *31*, 6–17.
33. Feged-Rivadeneira, A.; Del Cairo, C.; Vargas, W. Demographic and epidemic transitions in peri-urban areas of Colombia: a multilevel study of malaria in the Amazonian city of San José del Guaviare. *Environ. Urban.* **2019**, *31*, 325–348, doi:10.1177/0956247818808207.
34. Feged-Rivadeneira, A.; Evans, S. Ethnography of a parasite: A quantitative ethnographic observation of forest malaria in the Amazon basin. *Scand. J. Public Health* **2019**, *47*, 820–831, doi:10.1177/1403494818756561.
35. Bouma, A.; De Jong, M.; Kimman, T. Transmission of pseudorabies virus within pig populations is independent of the size of the population. *Prev. Vet. Med.* **1995**, *23*, 163–172.
36. Gandon, S.; Michalakis, Y. Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *J. Evol. Biol.* **2002**, *15*, 451–462.
37. Bronner, A.; Morignat, E.; Gay, E.; Vergne, T.; Fournié, G.; Pfeiffer, D.U.; Calavas, D. Iso-population partition: An innovative epidemiological approach to mapping and analyzing spatially aggregated data. *Prev. Vet. Med.* **2015**, *122*, 253–256.
38. Espinoza, L.M.C.; Nichols, C.; Adu-Sarkodie, Y.; Al-Emran, H.M.; Baker, S.; Clemens, J.D.; Dekker, D.M.; Eibach, D.; Krumkamp, R.; Boahen, K.; et al. Variations of invasive *Salmonella* infections by population size in Asante Akim North Municipal, Ghana. *Clin. Infect. Dis.* **2016**, *62*, S17–S22.
39. Brown, J.H.; Gillooly, J.F.; Allen, A.P.; Savage, V.M.; West, G.B. Toward a metabolic theory of ecology. *Ecology* **2004**, *85*, 1771–1789.
40. Gillooly, J.F.; Brown, J.H.; West, G.B.; Savage, V.M.; Charnov, E.L. Effects of size and temperature on metabolic rate. *science* **2001**, *293*, 2248–2251.
41. Tinbergen, N. On aims and methods of ethology. *Z. Für Tierpsychol.* **1963**, *20*, 410–433.
42. Stumpf, M.P.; Porter, M.A. Critical truths about power laws. *Science* **2012**, *335*, 665–666.
43. Harris, J.A.; Benedict, F.G. A biometric study of human basal metabolism. *Proc. Natl. Acad. Sci.* **1918**, *4*, 370–373.
44. Kleiber, M. Body size and metabolism. *ENE* **1932**, *1*, E9.
45. Gabaix, X.; Gopikrishnan, P.; Plerou, V.; Stanley, H.E. A theory of power-law



- distributions in financial market fluctuations. *Nature* **2003**, *423*, 267–270.
46. Stroud, P.D.; Sydoriak, S.J.; Riese, J.M.; Smith, J.P.; Mniszewski, S.M.; Romero, P.R. Semi-empirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing. *Math. Biosci.* **2006**, *203*, 301–318, doi:10.1016/j.mbs.2006.01.007.
  47. Massad, E.; Ma, S.; Chen, M.; Struchiner, C.J.; Stollenwerk, N.; Aguiar, M. Scale-free network of a dengue epidemic. *Appl. Math. Comput.* **2008**, *195*, 376–381, doi:10.1016/j.amc.2007.04.102.
  48. Singer, H.M. The COVID-19 pandemic: growth patterns, power law scaling, and saturation. *Phys. Biol.* **2020**.
  49. Parés-Ramos, I.K.; Álvarez-Berrios, N.L.; Aide, T.M. Mapping Urbanization Dynamics in Major Cities of Colombia, Ecuador, Perú, and Bolivia Using Night-Time Satellite Imagery. *Land* **2013**, *2*, 37–59, doi:10.3390/land2010037.
  50. Lai, W. *Fitting power law distributions to data*; 2016;
  51. Team, R.C. *R: A language and environment for statistical computing*; Vienna, Austria, 2013;
  52. Rader, E.; Wash, R. Influences on tag choices in del. icio. us. In Proceedings of the Proceedings of the 2008 ACM conference on Computer supported cooperative work; 2008; pp. 239–248.
  53. Knudson, A.; González-Casabianca, F.; Feged-Rivadeneira, A.; Pedreros, M.F.; Aponte, S.; Olaya, A.; Castillo, C.F.; Mancilla, E.; Piamba-Dorado, A.; Sanchez-Pedraza, R.; et al. Spatio-temporal dynamics of Plasmodium falciparum transmission within a spatial unit on the Colombian Pacific Coast. *Sci. Rep.* **2020**, *10*, 3756, doi:10.1038/s41598-020-60676-1.
  54. Ministerio de Salud y Protección Social *Plan Decenal de Salud Pública 2012-2021*; Imprenta Nacional de Colombia, 2012;
  55. Sandes, L.F.F.; Freitas, D.A.; de Souza, M.F.N.S.; Leite, K.B. de S. Primary health care for South-American indigenous peoples: an integrative review of the literature Atención primaria en salud a indígenas de América del Sur: revisión integrativa de la bibliografía. *Rev. Panam. Salud Publica Pan Am. J. Public Health* **2018**, *42*, e163, doi:10.26633/RPSP.2018.163.
  56. Noreña-Herrera, C.; Leyva-Flores, R.; Palacio-Mejía, L.S.; Duarte-Gómez, M.B. Inequidad en la utilización de servicios de salud reproductiva en Colombia en mujeres indígenas y afrodescendientes. *Cad. Saúde Pública* **2015**, *31*, 2635–2648, doi:10.1590/0102-311X00016515.
  57. Hurtado-Saa, T.; Rosas-Vargas, R.; Valdés-Cobos, A. Servicios de salud, discriminación y condición étnica/racial: un estudio de caso de la problemática en México y Colombia. *Ra Ximhai* **2013**, *9*, 135–151.
  58. Espinosa, A.E.; Aviléz, J.B.; Payares, A.U. Segregación residencial de afrodescendientes en Cartagena, Colombia. *Econ. Región* **2018**, *12*, 95–132.
  59. Lucumi, D.I.; LeBrón, A.M.W.; Schulz, A.J.; Mentz, G. Social patterning of cardiovascular and metabolic risk in Colombian adults. *Ethn. Health* **2017**, *22*, 389–401, doi:10.1080/13557858.2016.1244628.
  60. Lucumi, D.I.; Schulz, A.J.; Roux, A.V.D.; Grogan-Kaylor, A. Income inequality and high blood pressure in Colombia: a multilevel analysis. *Cad. Saúde Pública* **2017**, *33*, e00172316, doi:10.1590/0102-311x00172316.
  61. Bermúdez, A.; Flórez, O.; Bolaños, M.V.; Medina, J.J.; Salcedo-Cifuentes, M. Enteroparasitismo, higiene y saneamiento ambiental en menores de seis comunidades indígenas. Cali-Colombia. *Rev. Salud Pública* **2013**, *15*, 1–11.
  62. Quintero, L.; López, M.B.; Ramírez, H.; Castaño, J.C. Descripción de un brote

epidémico de malaria en una comunidad indígena asentada en la zona urbana de Armenia, Colombia, 2012. *Biomédica* **2015**, 35, 24–33.

63. Bettencourt, L.M. The origins of scaling in cities. *science* **2013**, 340, 1438–1441.