

EPIDEMIOLOGY

Scaling of mortality in 742 metropolitan areas of the Americas

Usama Bilal^{1,2*}, Caio P. de Castro^{3,4}, Tania Alfaro⁵, Tonatiuh Barrientos-Gutierrez⁶, Mauricio L. Barreto^{3,7}, Carlos M. Leveau^{8,9}, Kevin Martinez-Folgar^{1,2}, J. Jaime Miranda^{10,11}, Felipe Montes¹², Pricila Mullachery¹, Maria Fatima Pina^{13,14}, Daniel A. Rodriguez¹⁵, Gervasio F. dos Santos^{3,16}, Roberto F. S. Andrade^{3,4}, Ana V. Diez Roux^{1,2}

We explored how mortality scales with city population size using vital registration and population data from 742 cities in 10 Latin American countries and the United States. We found that more populated cities had lower mortality (sublinear scaling), driven by a sublinear pattern in U.S. cities, while Latin American cities had similar mortality across city sizes. Sexually transmitted infections and homicides showed higher rates in larger cities (superlinear scaling). Tuberculosis mortality behaved sublinearly in U.S. and Mexican cities and superlinearly in other Latin American cities. Other communicable, maternal, neonatal, and nutritional deaths, and deaths due to noncommunicable diseases were generally sublinear in the United States and linear or superlinear in Latin America. Our findings reveal distinct patterns across the Americas, suggesting no universal relation between city size and mortality, pointing to the importance of understanding the processes that explain heterogeneity in scaling behavior or mortality to further advance urban health policies.

INTRODUCTION

More than 50% of the global population lives in cities, and by 2050, this figure is expected to reach 70% (1). The process of urbanization has been especially intense in Latin America, a region that has undergone a rapid urbanization process in a brief period of time. North America also exhibits high levels of urbanization, making the Americas the most urbanized region in the world (2). Cities are complex systems of interacting agents that give rise to emergent phenomena, including levels and distributions of population health (3–6). Although much research has evidenced contrasting levels of health between urban and rural areas, findings have been mixed partly due to the heterogeneity of urban areas themselves (6). Our knowledge of what specific features of urban areas or cities affect population health, and of the processes through which they do so, is still limited (7).

Like biological organisms, as cities grow the complexity of their processes also grows (4, 5, 8). Population size can be conceptualized as an indicator of multiple socioenvironmental mechanisms linked to the agglomeration and intense interaction of people, which can result in advantages or disadvantages depending on the outcome in question. Some processes scale superlinearly with the population

size of cities, including socially generated outcomes such as wealth and crime, meaning that they occur at a higher rate in larger cities, potentially due to network effects and increased social interactions (4, 5). Other processes scale sublinearly, meaning that they increase at a slower rate than the population; for instance, infrastructure characteristics such as the road network length scale sublinearly (4, 5), leveraging economies of scale, such as the ability of roads to accommodate a higher number of users. Last, a number of outputs scale linearly, meaning that they occur at similar rates across the continuum of population size (4, 5). Exploring the specific processes that link health outcomes and population size in cities may shed light into underlying mechanisms that could be harnessed to improve population health in urban areas.

To date, a number of studies have examined how population size relates to health in cities (4, 9–19). Here, we report on a comprehensive examination of how mortality scales with city population size across the universe of 742 metropolitan areas of 100,000 residents or more in the United States and 10 Latin American countries. We advance and innovate on previous work in three ways. First, by investigating cities across 11 very different countries, we can draw more generalizable inferences regarding the urban scaling of mortality. Second, we provide an exhaustive examination of both fine-grained and coarse classifications of causes of death. Third, we conduct the analyses taking into consideration key epidemiologic and demographic aspects that may drive differences in mortality linked to city size, including age differentials across cities and coding differences across countries. These results provide a comprehensive understanding of how health indicators scale with city size. Knowledge of the fundamental drivers of health in cities is critical to identifying the strategies that can be used to promote health in urban areas.

RESULTS

Urban scaling of all-cause mortality

We first explored the scaling properties of all-cause and cause-specific mortality in 742 cities pooling data from 2010 to 2016. We found

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

¹Urban Health Collaborative, Drexel Dornsife School of Public Health, Philadelphia, PA, USA. ²Department of Epidemiology and Biostatistics, Drexel Dornsife School of Public Health, Philadelphia, PA, USA. ³Center for Data and Knowledge Integration for Health (CIDACS), Fiocruz, Salvador, Bahia, Brazil. ⁴Institute of Physics, Federal University of Bahia, Salvador, Bahia, Brazil. ⁵Escuela de Salud Pública, Universidad de Chile, Santiago de Chile, Chile. ⁶Instituto Nacional de Salud Pública, Cuernavaca, Mexico. ⁷Institute of Collective Health, Federal University of Bahia, Salvador, Bahia, Brazil. ⁸Instituto de Producción, Economía y Trabajo, Universidad Nacional de Lanús, Buenos Aires, Argentina. ⁹Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina. ¹⁰CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru. ¹¹School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru. ¹²Department of Industrial Engineering, Universidad de los Andes, Bogotá, Colombia. ¹³Institute for Information and Communication on Health—ICICT/FIOCRUZ, Rio de Janeiro, Brazil. ¹⁴i3S—Instituto de Investigação e Inovação em Saúde, Porto, Portugal. ¹⁵Department of City and Regional Planning, University of California, Berkeley, Berkeley, CA, USA. ¹⁶Economics Faculty, Federal University of Bahia, Salvador, Bahia, Brazil.

*Corresponding author. Email: ubilal@drexel.edu

that all-cause mortality scaled sublinearly, with $\beta = 0.94$ [95% confidence interval (CI), 0.92 to 0.96] (Fig. 1, Table 1, and fig. S1). In other words, more populated cities have a relatively lower mortality rate compared to smaller cities. After adjusting for age and country, the sublinear scaling coefficient was minimally attenuated ($\beta = 0.97$; 95% CI, 0.96 to 0.97), indicating that, even after considering the different distribution of ages across cities among the entire region of the Americas, larger cities had lower all-cause mortality. We also found that this pattern differed by region. Specifically, after adjusting

for age and country, the scaling coefficient for all-cause mortality was sublinear in U.S. cities ($\beta = 0.94$; 95% CI, 0.93 to 0.95) but linear in Latin American cities ($\beta = 1.00$; 95% CI, 0.99 to 1.01). Therefore, the health advantage of living in larger cities was present in U.S. cities and was absent in Latin American cities. The scaling pattern for Latin American cities was similar in cities of the two largest countries, Mexico ($\beta = 1.01$; 95% CI, 0.98 to 1.03) and Brazil ($\beta = 1.00$; 95% CI, 0.99 to 1.01), and other Latin American countries ($\beta = 0.99$; 95% CI, 0.97 to 1.01).

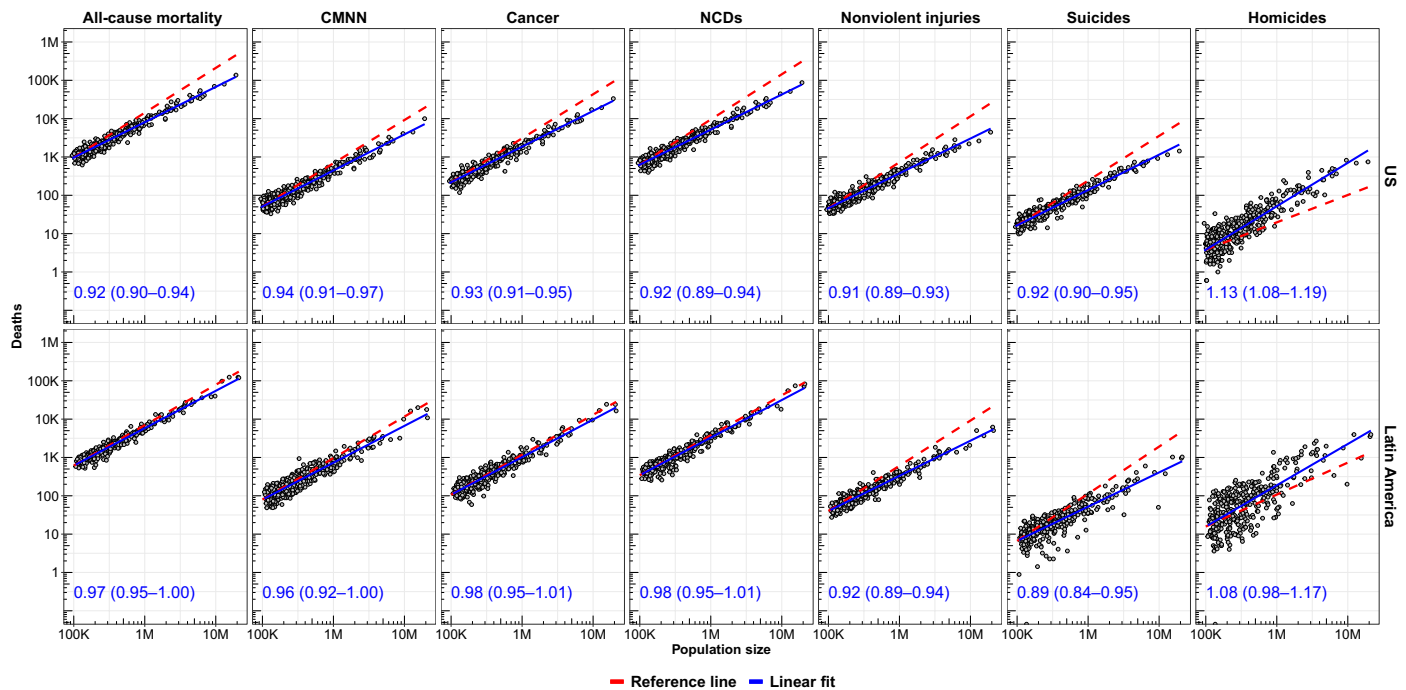


Fig. 1. Scaling of all-cause and cause-specific mortality relative to city population size in U.S. and Latin American cities. Solid blue lines are linear fits of log(deaths) on log(population); red dashed lines are reference lines ($\beta = 1$). Coefficients (95% CI) are unadjusted coefficients of log(deaths) on log(population), stratified by region. CMNN, communicable, maternal, neonatal, and nutritional conditions; NCDs, noncommunicable diseases.

Table 1. Scaling coefficients (β , 95% CI) by cause of death for all U.S. and Latin American cities. CMNN, communicable, maternal, neonatal, and nutritional diseases; CVD/NCDs, cardiovascular disease and other noncommunicable diseases.

Cause/group	Unadjusted	Adjusted*	United States [†]	Latin America [†]	BR [†]	MX [†]	Latin America (no BR/MX) [†]
All-cause mortality	0.94 (0.92–0.96)	0.97 (0.96–0.97)	0.94 (0.93–0.95)	1.00 (0.99–1.01)	1.00 (0.99–1.01)	1.01 (0.98–1.03)	0.99 (0.97–1.01)
CMNN	0.96 (0.93–0.99)	0.97 (0.95–0.99)	0.95 (0.92–0.97)	1.01 (0.98–1.03)	1.01 (0.98–1.04)	0.99 (0.95–1.03)	1.02 (0.98–1.07)
Cancer	0.94 (0.91–0.97)	0.98 (0.97–0.99)	0.95 (0.94–0.97)	1.01 (1.00–1.03)	1.01 (1.00–1.03)	1.01 (0.98–1.04)	1.00 (0.97–1.03)
CVD/NCDs	0.93 (0.91–0.96)	0.96 (0.95–0.97)	0.94 (0.92–0.95)	1.00 (0.99–1.01)	0.99 (0.98–1.01)	1.02 (0.99–1.04)	0.99 (0.96–1.01)
Nonviolent injuries	0.91 (0.89–0.93)	0.93 (0.91–0.94)	0.92 (0.90–0.94)	0.93 (0.90–0.95)	0.93 (0.90–0.97)	0.94 (0.90–0.99)	0.90 (0.85–0.95)
Suicides	0.88 (0.84–0.93)	0.92 (0.89–0.94)	0.94 (0.92–0.97)	0.88 (0.84–0.92)	0.88 (0.83–0.93)	0.91 (0.82–1.00)	0.87 (0.79–0.95)
Homicides	1.14 (1.07–1.22)	1.12 (1.07–1.16)	1.12 (1.07–1.18)	1.10 (1.04–1.17)	1.17 (1.09–1.25)	0.97 (0.80–1.13)	1.01 (0.91–1.12)

*Adjusted model is adjusted by age structure and country. [†]Stratified models are run only on the indicated sample (e.g., Latin America is ran with all Latin American cities), all adjusted by age structure and country (where relevant). For Latin American cities, the main analysis includes the 366 cities in 10 countries, while BR includes Brazilian cities ($n = 152$), MX includes Mexican cities only ($n = 92$), and “no BR/MX” includes all Latin American cities except for those in BR and MX ($n = 122$).

Urban scaling of large groupings of causes of death

Next, to determine whether these overall scaling patterns of all-cause mortality obscure differing patterns by cause, we explored whether the sublinear (for the United States) and linear (for Latin America) scaling of all-cause mortality applied to the six large groupings of causes of death (Table 1). The three large disease groupings [CMNN (communicable, maternal, neonatal, and nutritional condition), cancer, and NCDs (noncommunicable diseases)] had sublinear scaling in U.S. cities ($\beta = 0.95, 0.95,$ and 0.94) and linear scaling in Latin American cities ($\beta = 1.00, 1.01,$ and 1.01). In contrast, nonviolent injuries ($\beta = 0.92$ in the United States and 0.93 in Latin America) and suicides ($\beta = 0.94$ and 0.88) had a sublinear pattern in both regions, while homicides had a superlinear pattern ($\beta = 1.12$ and 1.10). The sublinear pattern for nonviolent injuries and suicides for Latin America was similar for Brazilian, Mexican, and all other Latin American cities ($\beta = 0.93, 0.94,$ and 0.90 for nonviolent injuries and $\beta = 0.88, 0.91,$ and 0.87 for suicides). However, the superlinear pattern for homicides varied widely, being superlinear only for Brazilian cities ($\beta = 1.17$) and linear for Mexican and other Latin American cities ($\beta = 0.97$ and 1.01).

Of note, model fit for most outcomes by region was very good, with all R^2 above 80% and most above 90%, with the exception of homicides in Mexico ($R^2 = 65%$, compared to 81% in the United States, 83% for all cities in Latin America, 87% for Brazil, and 89% for all other Latin American cities; see table S2), indicating a higher variability in homicide rates across cities in Mexico after accounting for city size, suggesting stronger place-based (city-specific) effects in Mexico. Later in this section, we will describe a more comprehensive exploration of the relationship between this variability in outcomes and scaling behaviors. Figure 2 summarizes the ordering of scaling coefficients by model, with causes of death that are more superlinear on top. We note that, irrespectively of the superlinear, linear, or sublinear character of the results, the ordering of the six large groupings of causes of death was similar for U.S. and Latin American cities as a whole. This means that, while the specific

scaling coefficients varied by region, the ordering of causes from most to least superlinear was mostly conserved. The only exception was suicides, which were the most sublinear cause in Latin America and the fourth most sublinear cause in U.S. cities. Brazilian cities followed a very similar pattern to Latin American cities as a whole. In contrast, in Mexican cities, homicides ranked only fifth in terms of superlinearity (as compared to first in U.S. cities, Latin American cities as a whole, and Brazilian cities). NCDs (which were linear for the Latin American region as a whole) were weakly superlinear in Mexico and ranked first in terms of superlinearity (as compared to fifth, fourth, and fifth in U.S. cities, Latin American cities as a whole, and Brazilian cities, respectively).

Urban scaling of detailed causes of death

Given the heterogeneity in causes of death within these large six groupings, we also conducted a more fine-grained analysis focused on specific death causes (see Figs. 3 and 4 and tables S3 to S5 for results).

Communicable, maternal, neonatal, and nutritional conditions

Among the four communicable disease conditions [sexually transmitted diseases (STDs) and HIV/AIDS, tuberculosis, respiratory infections, and all other infectious diseases], we found different patterns by region and cause. STDs and HIV/AIDS had a superlinear behavior for U.S. cities, Latin American cities as a whole, Brazilian cities, and all other Latin American cities ($\beta = 1.24, 1.07, 1.07,$ and $1.17,$ respectively), but they were sublinear in Mexico ($\beta = 0.90$). STDs and HIV/AIDS were actually the most superlinear cause of death in U.S. cities and in Latin American cities excluding Brazil and Mexico, but they were the second most sublinear cause in Mexico.

On the other hand, tuberculosis had a very different pattern in U.S. and Latin American cities (Fig. 4A): It was the second most superlinear cause in Latin American cities but the second most sublinear cause in U.S. cities ($\beta = 1.10$ and 0.88 for Latin American and

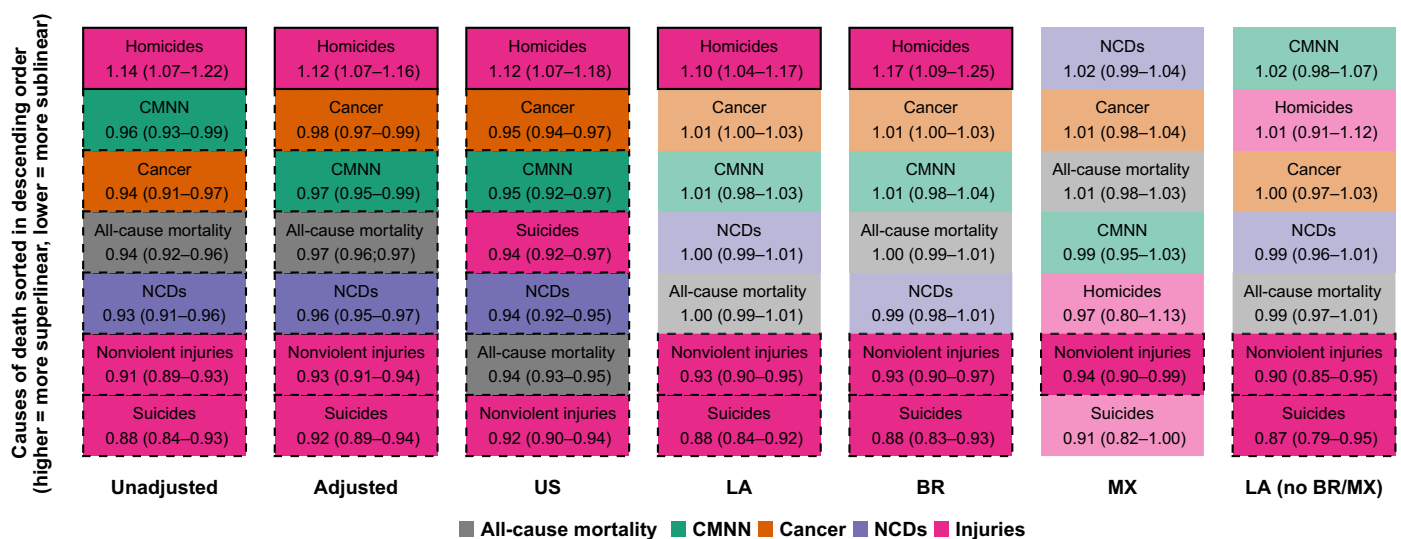


Fig. 2. Large groupings of causes of death sorted by scaling coefficient. Fully colored cells indicate a statistically significant superlinear or sublinear pattern; cells with a solid outline indicate a superlinear pattern; cells with a dashed outline indicate a sublinear pattern; non-fully colored cells with no outline indicate a coefficient whose 95% CI crosses the null of linearity.

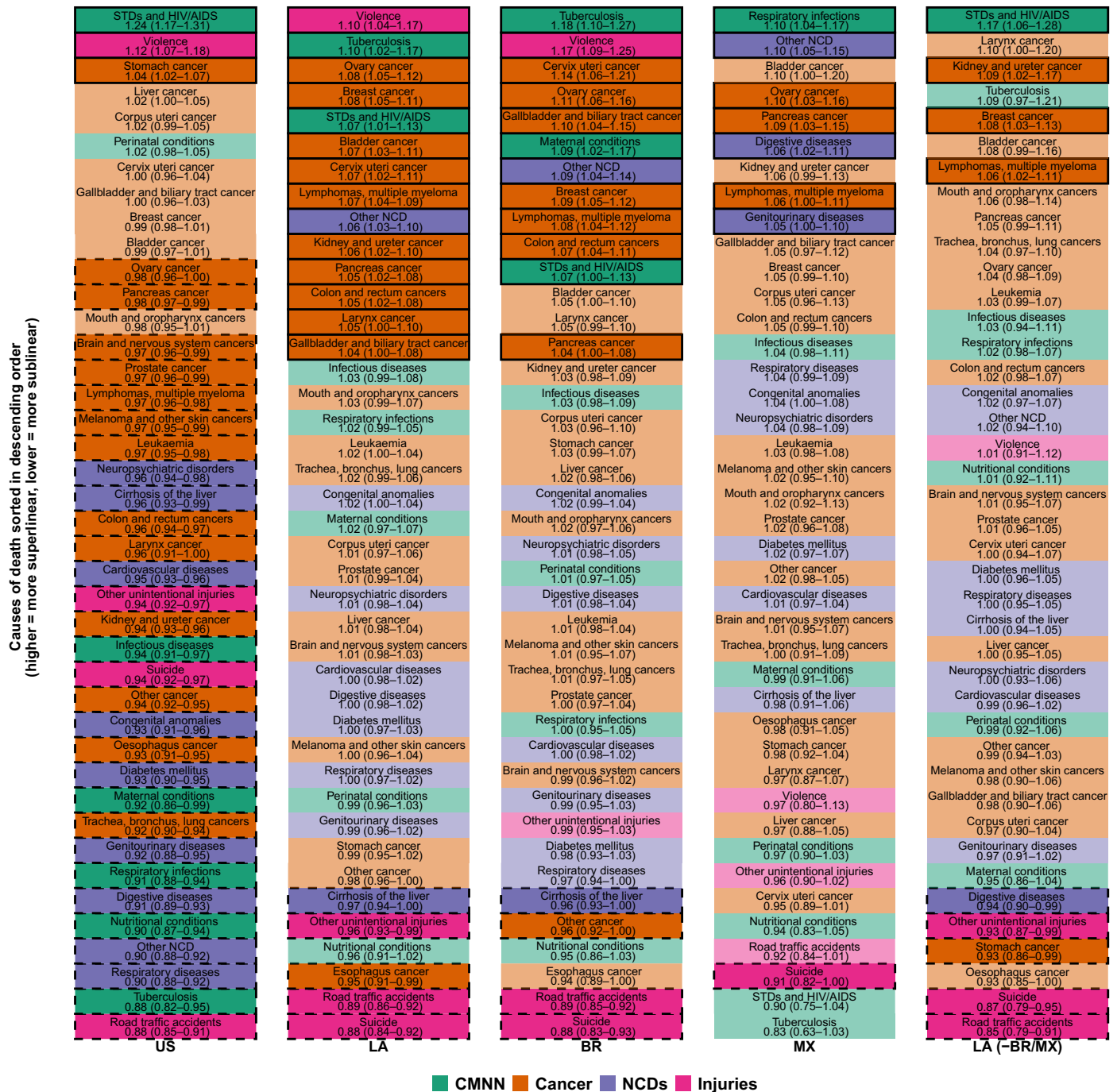


Fig. 3. Causes of death sorted by scaling exponent and region. Fully colored cells indicate a statistically significant superlinear or sublinear pattern; cells with a solid outline indicate a superlinear pattern; cells with a dashed outline indicate a sublinear pattern; non-fully colored cells with no outline indicate a coefficient whose 95% CI crosses the null of linearity.

U.S. cities, respectively). The pattern in Latin American cities was driven by Brazil and all other Latin American cities ($\beta = 1.18$ and 1.09, respectively), as tuberculosis was the most sublinear cause in Mexican cities ($\beta = 0.83$). In general, all other communicable disease conditions, along with maternal, neonatal, and nutritional conditions, had a sublinear behavior in U.S. cities (ranking among the most sublinear causes of death) and a linear behavior in Latin

American cities, with two exceptions: Maternal conditions were superlinear ($\beta = 1.09$) in Brazilian cities, and respiratory infections were superlinear in Mexican cities ($\beta = 1.10$).

Noncommunicable diseases

We observed that mortality from most cancers showed a sublinear (13 of the 21 subcategories) or linear (7 of 21) pattern in U.S. cities.

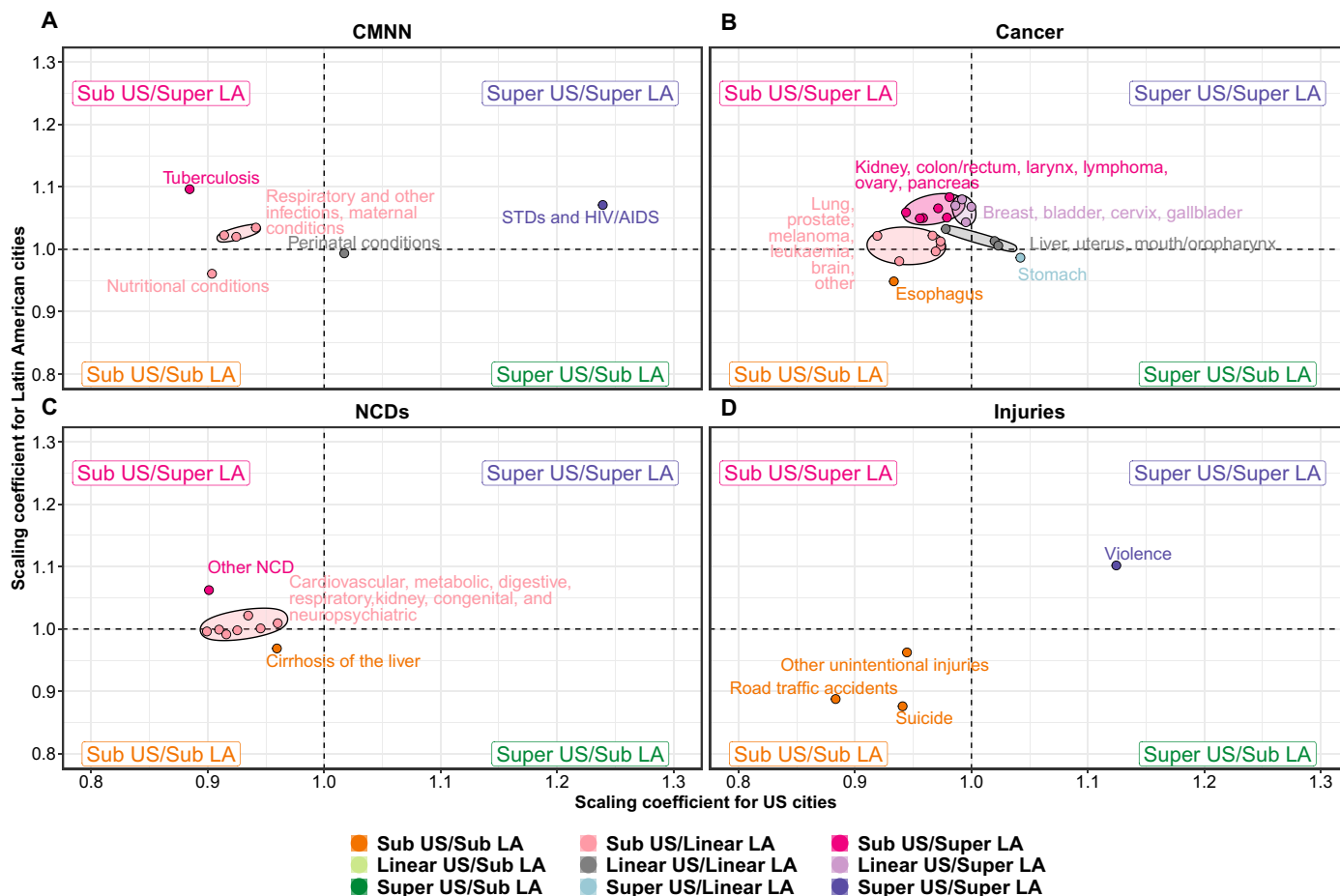


Fig. 4. Comparison of scaling coefficients in U.S. versus Latin American cities by cause. Coefficients come from a model adjusted for country and age distribution and stratified by region (Eq. 3).

In contrast, cancer deaths showed a linear (10 of 21) or superlinear (10 of 21) pattern in Latin American cities. Specifically, and as seen in Fig. 4B, we found a group of cancer types that showed sublinear patterns in the United States and superlinear patterns in Latin America: kidney, colon/rectum, larynx, lymphomas and multiple myelomas, ovary, and pancreas; a second group of cancer types showed a sublinear pattern in the United States but a linear pattern in Latin America (lung and trachea, prostate, melanoma, leukemia, and brain), while esophagus cancer showed a sublinear pattern in both regions. Stomach cancer showed a superlinear pattern in the United States and a linear pattern in Latin America, while liver, uterus, and mouth/oropharynx cancers had a linear pattern in both regions. Last, breast, bladder, cervix, and gallbladder cancer had a superlinear pattern in Latin America and a linear pattern in the United States.

Cardiovascular and other NCDs were clustered into three groups (Fig. 4C). The first group included cirrhosis of the liver, which showed a sublinear pattern in both regions. However, its rank with respect to other NCDs varied by region (Fig. 3), as cirrhosis was the second least sublinear NCD in U.S. cities while being the most sublinear NCD in Latin American cities as a whole, and in Brazilian and Mexican cities. A second group included several other NCDs

(oral diseases, musculoskeletal conditions, skin diseases, and sense organ diseases), which had a clear superlinear pattern (and were the most superlinear NCD) in Latin American, Brazilian, and Mexican cities ($\beta = 1.06, 1.09, 1.10$, respectively) while having a clear sublinear pattern in U.S. cities ($\beta = 0.90$), ranking as the second most sublinear NCD. The last group, including seven of the nine NCDs (cardiovascular, neuropsychiatric, congenital anomalies, diabetes, genitourinary, digestive, and respiratory diseases), were sublinear in U.S. cities and linear in Latin American cities as a whole. However, the ordering of these causes was similar by region, as neuropsychiatric and cardiovascular diseases and congenital anomalies were among the least sublinear NCDs in both regions, while genitourinary, digestive, and respiratory diseases were among the most sublinear NCDs in both regions (Fig. 3). Again, as with other causes of death, these patterns held for Latin American cities as a whole and Brazilian cities, but differed for cities of Mexico, where digestive, genitourinary, and respiratory diseases were among the most superlinear NCDs.

Injuries

In the case of injury deaths, patterns were mostly consistent between U.S. and Latin American cities (Figs. 3 and 4D). Suicides, road traffic deaths, and other unintentional injuries showed a sublinear pattern

in both regions, with a similar magnitude ($\beta = 0.94$ and 0.88 for suicides, $\beta = 0.88$ and 0.89 for road traffic deaths, and $\beta = 0.94$ and 0.96 for other unintentional injuries in U.S. and Latin American cities, respectively). However, their ranking with respect to other causes varied, as they represented some of the most sublinear causes in Latin American cities (including the group of Brazilian cities, Mexican cities, and all other cities excluding Brazil and Mexico), while only road traffic accidents had a similar behavior in the United States, as suicides and other intentional injuries (which include falls and drug overdoses) had an average rank (but were still sublinear). On the other hand, homicides followed a superlinear scaling pattern in both regions ($\beta = 1.12$ and 1.10 for U.S. and Latin American cities, respectively). As described above, the only exception to these patterns was for homicides in Mexican and other Latin American cities, which followed a linear pattern [$\beta = 0.97$ (95% CI, 0.80 to 1.13) and $\beta = 1.01$ (95% CI, 0.91 to 1.12) for Mexican and other Latin American cities, respectively]. Homicides ranked among the most sublinear causes of death for Mexican cities (Fig. 3).

Secondary analyses

We also explored the association between the levels and variance of each outcome and their respective scaling coefficient (Fig. 5). We

found a negative correlation between scaling exponents and levels (intercepts in the models) ($\rho = -0.58$ and -0.67 for U.S. and Latin American cities, respectively), indicating that causes of death that are more frequent (higher intercepts) tend to be more sublinear, while causes of death that are less frequent (lower intercepts) tend to be superlinear, similarly across both regions (Fig. 5A). We found a weak positive correlation between scaling exponents and the square root of the mean squared error ($\rho = 0.40$ and 0.31 for U.S. and Latin American cities, respectively), indicating that causes of death that vary more tend to be more superlinear, while those that have lower variability tend to be more sublinear (Fig. 5B). Sensitivity analyses testing two alternative definitions for U.S. cities, commuting zones and an ad hoc definition based on which counties overlap with the urban area of the city, and adding an indicator to the main model indicating whether the city was the largest in its country all rendered similar results as our main analyses (figs. S2 and S3).

DISCUSSION

Our analysis of 366 Latin American and 376 U.S. cities (encompassing over 560 million residents) revealed a heterogeneous scaling

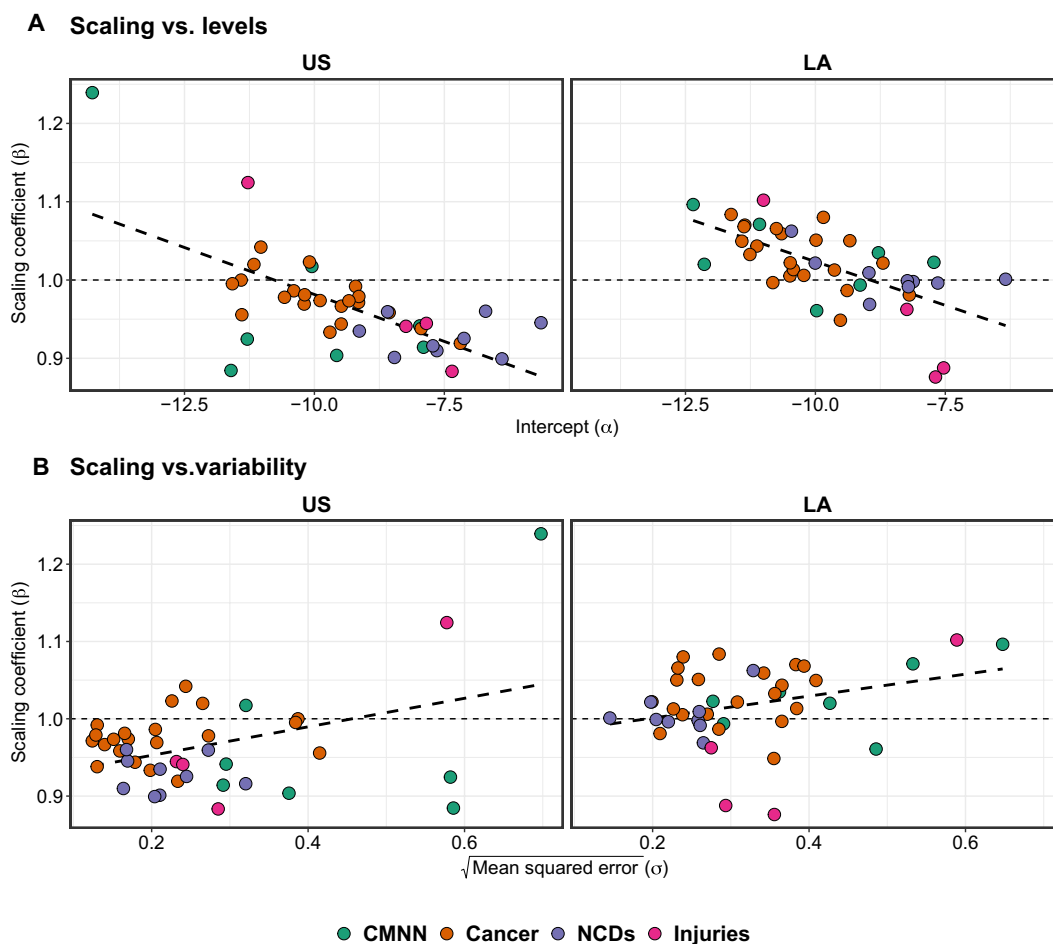


Fig. 5. Correlation between scaling exponents and intercepts (levels) and standard deviations (variability) for each cause of death, by region. Coefficients come from the model in Eq. 3, stratified by region.

landscape of mortality across the continent. While in the United States all-cause mortality was relatively lower in larger cities and did not differ by size in Latin American cities, the relation between city size and mortality differed by cause of death and region. Nonviolent injury deaths were sublinear in both regions, whereas homicides and deaths due to STDs/AIDS were superlinear in both regions. NCDs and cancer were generally sublinear (or linear for some cancers) in the United States and mostly linear or superlinear in Latin America. Deaths from CMNN (other than AIDS/STDs) were sublinear in the United States and linear in Latin America.

To date, a number of studies have examined the scaling properties of population health outcomes, most of which focus on single conditions, including STDs and injuries, or cities from a single country. Studies in the United States and Brazil have shown that both HIV/AIDS and STDs scale superlinearly (4, 10, 11, 13), consistent with our findings in both the United States and Latin America. In the case of injuries, several studies in the United States and Latin America are consistent with our findings of an overall superlinear scaling pattern of homicides, sublinear scaling pattern of suicides, and sublinear or linear scaling pattern for road traffic injuries (4, 14, 16, 20–24). Rocha *et al.* (12) examined a wider variety of outcomes in U.S. counties, as well as Swedish and Brazilian municipalities, but did not do so in metropolitan areas. Choi *et al.* (17) also examined the scaling of mortality by four large NCD groupings in U.S. counties, finding a superlinear pattern for endocrine and metabolic diseases and a sublinear pattern for other NCDs when considering the largest counties.

Some causes of death show a consistent scaling pattern across regions

The consistency of findings regarding sublinear scaling of nonviolent injury deaths and superlinear scaling of homicides and deaths due to STDs/AIDS across both regions and a diverse set of cities is notable in terms of its consistency, suggesting that some underlying common dynamics may play a role, irrespective of cross-country differences. For example, the sublinear scaling of nonviolent deaths may be because larger cities may have attributes such as more or better public transportation, lower traffic speeds, or sidewalk infrastructure (25) that may be linked to lower rates of traffic-related deaths, an important contributor to nonviolent deaths. The sublinear pattern of pedestrian deaths, a component of road traffic deaths, has been reported before in the United States (16) and Brazil (20), linked to a shorter road traffic network in larger cities (26) and higher traffic congestion leading to lower speeds (16).

The superlinear scaling of homicides may be linked to differences in social conditions and inequality by city size. Larger cities are often characterized by large inequalities and social exclusion (27), originating from both segregation and self-sorting (28–31) and differential migration to/from cities (30, 32, 33). On the other hand, larger cities also display wider diversity of economic activities, which may result in improved social and economic outcomes (32, 34). Combinations of some of these factors, along with increased social interactions (35), may be linked to higher homicide rates in larger cities, consistent with previous research in the United States and Latin America (14). Consistent with these findings, other studies have documented higher rates of crime in larger cities (4, 14), including homicides. The superlinearity of HIV/AIDS and STDs, reported in several previous studies (4, 10, 11), is likely to emerge from a combination of increased social contacts in larger cities and social, economic, and behavioral differences between cities of different sizes, which leads

to increased incidence of both syphilis, gonorrhea, and chlamydia (10, 11) and HIV/AIDS (4, 13).

Other infectious diseases, and maternal, neonatal, and nutritional conditions display different patterns by region

Deaths from other CMNN conditions were sublinear in the United States and linear or even superlinear in Latin America. This highlights that the distribution of factors driving these conditions is different in both regions. For example, tuberculosis is strongly sublinear in the United States and superlinear in Latin America, especially in Brazilian and other Latin American cities excluding Mexico. Large cities in low and middle income countries show very wide heterogeneity (36) in the conditions that give rise to tuberculosis, including poverty and overcrowding, and may also differ in access to high-quality health care services (37). In contrast, cities in high-income countries may have better infrastructure. This was shown in a recent study that highlighted that deaths due to conditions for which the health care system can prevent new cases, which include tuberculosis, were especially high in large cities of Latin America (38). Migration may be another factor behind the sublinear pattern of tuberculosis mortality in the United States. In 2011, 62% of reported tuberculosis cases in the United States occurred among foreign-born people, mainly Hispanic and Asian individuals (39). Barriers to accessing care associated with immigration status can be one of the main factors preventing adherence to treatment, especially among non-English-speaking migrants (40). This phenomenon may be more frequent in small and medium-sized cities, given the geographic dispersion of Hispanic immigration (41).

The scaling of NCDs varies widely by specific cause

The different scaling of cancer and NCDs in Latin America (mostly linear) and United States (mostly sublinear) may be related to differential associations of chronic disease risks factors with city size in both regions, particularly considering that most of Latin America has swiftly transitioned toward a predominance of urban areas and the accompanying epidemiological and nutritional transitions that are still in progress (42–44). Smoking, excessive alcohol consumption, and obesity are more prevalent in smaller compared to larger cities of the United States (45). To our knowledge, very few studies have compared NCD risk factor levels across the spectrum of urbanization in Latin America. In an analysis of NCD risk factors by province-level urbanicity in Argentina, Rodríguez López *et al.* (46) found that more urbanized provinces had a higher prevalence of smoking but a lower prevalence of hypertension, but these patterns held only for women. Rural to urban comparisons have been more frequent, finding a consistently higher prevalence of NCD risk factors in urban compared to rural areas of Latin America (47–49). Various environmental and policy features linked to the prevalence of these risks factors [including factors such as air pollution, access to healthier and processed foods, walkability, and green spaces (50, 51)] may be strongly linked to city size in the United States and not so strongly linked to city size in Latin America, as has been observed in a comparison of U.S. and Indian and Chinese cities, which show very different scaling patterns for air pollution (52). However, an analysis of our same sample of 366 Latin American cities has shown that larger cities of Latin America have higher levels of air pollution, as measured by PM_{2.5} levels (53). Larger cities of both regions may also have higher health care capacity (54–58), although a recent study in

Latin American cities showed that deaths due to conditions for which the health care system should prevent cases are more frequent in larger cities, while deaths due to chronic conditions requiring substantial care are lower in larger cities (38).

The mostly sublinear scaling observed for cancers in the United States could also reflect better cancer survival in larger cities (59) as cancer mortality rates are especially sensitive to survival differences. Diagnostic capabilities may also be associated with city size, and these associations may be different in the United States and Latin America. We explored the scaling properties of ill-defined deaths, a marker of quality of coding of causes of death and potentially to improved diagnostic capabilities, and found them to be similarly sublinear in both regions ($\beta = 0.95$).

Implications

Our findings have implications for understanding the phenomenon of urban scaling and for prevailing urban scaling theories, which see the phenomenon as driven by common universal mechanisms rather than place-specific effects. The inclusion of cities in 10 different countries of Latin America and the United States and the study of multiple causes of death allowed us to explore heterogeneities by region and cause of death. Although we found some commonalities, we also observed important differences in the scaling behavior of specific causes of death and, in some cases, in the scaling behavior of a given cause of death across regions or countries. These differences were reflected in different values of scaling coefficients, their rankings within country or region, or both. We posit three possible explanations for these patterns.

First, and as outlined in the previous sections, the associations of health-relevant city characteristics including social, environmental, health care-related, and behavioral factors with city size may differ by domain and by region, and these features may relate differently to different causes of death. In other words, because different processes link city size to different health outcomes and the links between the drivers of these processes and city size may additionally vary by region, it is not surprising that the scaling of mortality varies by specific cause and region.

Second, and as proposed by Arcaute *et al.* (60) and Pumain (61), scaling patterns may be affected by path dependencies (historical contingencies) that influence specific features of cities regardless of their size (e.g., the role of San Francisco as a technological hub, despite its relatively small size compared to other cities). As a result, the context in which each city grew to its current size may affect the consequences of that growth and, therefore, the currently observed relationship between city size and mortality. Relatedly, Jedwab and Vollrath (62) have proposed a model of megacity growth, where they posit that current features of megacities vary depending on whether these cities grew before or after the transition to the current low mortality regime.

Third, as suggested by Gomez-Lievano *et al.* (32), scaling may also be a function of the nature of multifactorial causal processes at work. According to Gomez-Lievano *et al.* (32), outcomes that require the presence of multiple different factors (and that are therefore less common and have more variability) are more likely to occur in larger cities (behave superlinearly) because these multiple factors are more likely to co-occur in larger cities. Consistent with this hypothesis, we found that less common and more variable causes of death tended to be more superlinear. However, these associations do not necessarily prove the theory as, for example, it is not

immediately obvious that homicide and STD/AIDS deaths (strongly superlinear in our data) require more factors to co-occur than nonviolent injury deaths (which were strongly sublinear).

The stage of the epidemiological transition (63) of each country could affect scaling of health through each of the three explanations for scaling behaviors that we describe: through context and outcome-specific processes, through their histories and path dependencies, and by altering the relative prevalence of outcomes that require multiple factors to occur. The heterogeneity in stages of the epidemiologic transition reflected in countries we analyzed (64–66) could thus also explain heterogeneities in scaling coefficients across regions and countries.

Regardless of the processes that drive the scaling phenomena we observed, these relations have implications for public health interventions and urban policy. Knowing whether certain health outcomes scale with city size, and how, may allow for more precise resource allocation. For example, if two diseases show opposing scaling patterns, resources to prevent one with superlinear scaling could be focused on larger cities, while resources to prevent one with sublinear scaling could be focused on smaller cities. Moreover, a greater understanding of the drivers of the scaling phenomena could provide insights on whether there is an optimal city size (an important consideration for urban policy), a topic that has been extensively studied from a productivity perspective (67, 68), but rarely analyzed for population health optimization (69).

Caveats and limitations

There are some concerns regarding the quality of mortality data obtained from vital registration. First, there is a known underregistration of death counts that varies by country and subnationally (65, 70), which we addressed by applying an ensemble of state-of-the-art demographic methods to address this at the city level for Latin American cities, as reported elsewhere (65, 71). Second, a number of deaths are coded using ill-defined causes [International Classification of Diseases International Classification of Diseases (10th version) (ICD-10) codes R00 to R94 and R96 to R99], which we addressed by redistributing ill-defined deaths into more specific categories, based on age, sex, country, and year, as done in other studies (65). Ill-defined diseases scaled sublinearly in both U.S. cities and Latin American cities ($\beta = 0.95$), indicating improved coding of causes of death in larger cities, as compared to smaller cities, although the CIs crossed linearity in both cases (95% CI, 0.89 to 1.02 and 0.87 to 1.03 in U.S. and Latin American cities, respectively). Moreover, the coding of causes of death, and our grouping of causes of death, also has limitations. For example, our categorization of deaths due to communicable, maternal, neonatal, and nutritional conditions proved to be highly heterogeneous, as an analysis with more detailed causes of death showed very different scaling coefficients. This classification is rooted in the idea of epidemiologic transitions (63), which has been criticized (72) for, among other things, ignoring the potential emergence of new infectious diseases [such as HIV/AIDS or, more recently, coronavirus disease 2019 (COVID-19)]. Improvements in the categorization of causes of death may provide for more consistent results of the scaling patterns of deaths by cause.

A third caveat to our findings is related to the sensitivity of urban scaling properties to the definition of what constitutes a city and its boundaries (60). However, our sensitivity analysis exploring three alternative definitions for U.S. cities and our addition of an adjustment covariate for the largest city in each country [to control for the

“dragon-king” cities described by Arcaute *et al.* (60)] showed no changes to scaling patterns. Fourth, we are analyzing data for 11 countries, with their own regional differences in their urban systems. Historical factors have influenced the population and characteristics of cities in each country, resulting in potential challenges comparing cities of the same size across different countries. For example, job opportunities and the most advanced health care facilities in each country tend to concentrate in their largest cities (54–58). However, the size of these largest cities varies widely by country: The largest cities in the United States, Brazil, Mexico, and Argentina are all above 15 million, the largest cities in Chile, Colombia, and Peru have between 6 and 10 million residents, while the largest cities in Costa Rica, Guatemala, El Salvador, and Panama are all below 3.2 million residents. Therefore, if the likelihood of death increases (or decreases) with city size, as size approaches the largest city in a country, cities in smaller countries may not be comparable to similarly sized cities in larger countries. We controlled our main analysis for country, by adding a dummy covariate specific to each country, but we cannot rule out regional differences in each country that may not be captured by this adjustment. Last, our results are drawn from a cross-sectional sample and should not be interpreted directly as a measure of how much a city should grow (or shrink) to achieve that optimal size. That assertion assumes that cross-sectional associations reflect longitudinal associations and also requires the assumption of ergodicity, or lack of path dependence (61), which assumes that a given outcome depends only on the current state of the city, with no consideration to the path it took to get there, which may (73) or may not hold for cities (61, 74). Future studies should leverage longitudinal data that allow for the exploration of how city features change over time.

As the world continues urbanizing and as the COVID-19 pandemic brings scrutiny to the assumption that urban living translates into population-wide benefits, including better health profiles, it is especially important to deepen our understanding of how city size is related to health. Sublinear scaling of some causes of death indicates that larger cities can benefit from efficiency in services, from educational and job opportunities, and from environments and policies that may promote health. In contrast, superlinear scaling highlights the potentially negative correlates of larger cities, such as crowding, pollution, violence, and inequality.

Our results characterizing a comprehensive set of mortality patterns across a wide range of cities in the entire region of the Americas support the idea that there is no unique relation between city size and different health conditions, and that the application of urban scaling theories, which see the phenomenon as driven by common universal mechanisms (4, 5), to health outcomes, may need to be adapted to explain this heterogeneity. The inclusion of a large number of Latin American cities in our study advances the field by exploring the phenomenon of urban scaling in countries at varying income levels. Our findings raise questions about how city attributes can be leveraged to be health promoting while minimizing any adverse consequences. Understanding the processes that explain the heterogeneity in scaling behavior of mortality that we observed could be useful for health and urban policy.

It is possible that the ability of cities to maximize beneficial health, social, physical, and service environments declines above a certain size. Identifying the presence of a tipping point and understanding the historically and socially situated factors that drive its location are crucial for planning purposes and for public health

preparedness and responsiveness. This task requires considering health, energy efficiency, environmental, and economic impacts, much in line with the global interconnected targets set forth as the Sustainable Development Goals. An optimal city size could be an important universal social goal, minimizing adverse environmental impacts, maximizing health benefits, and sustaining an increased level of creativity and innovation that has long been a major characteristic of cities and the reason for its success to become the preferred shelter for the human population.

MATERIALS AND METHODS

Study setting

We used data on 366 Latin American and 376 U.S. cities. Latin American cities were defined as urban agglomerations of administrative units (municipios, comunas, distritos, partidos, etc.) that overlapped with the urban extent of the city, in 10 countries (65, 71): Argentina, Brazil, Chile, Colombia, Costa Rica, El Salvador, Mexico, Peru, and Panama. U.S. cities were defined as core-based statistical areas or the agglomeration of counties adjacent (and connected through commuting patterns) or part of a core area with at least 10,000 people. To make analyses comparable across both settings, we restricted our analysis to Latin American and U.S. cities with more than 100,000 people in 2010. We pooled data for the 2012–2016 period except for El Salvador (2010–2014), due to data availability.

Data sources

Data for Latin America were obtained from the Salud Urbana en America Latina (SALURBAL) study, which has compiled and harmonized vital registration and other health data (65, 71). Data for the United States were obtained from the National Vital Statistics System (75) and the Census Bureau. In all cases, we obtained all mortality records for the time frame of the study georeferenced to the county or county-equivalent level, with data on cause of death. We also obtained intercensal population estimations or postcensal projections by county or county-equivalent and age.

Variables

The main exposure investigated was the average yearly city population size in the period of the study, henceforth referred as city size. The main outcomes investigated are average yearly mortality counts by cause by city of residence. We classified causes of death based on the categories of the Global Health Estimates classification (76). Causes of death were first divided into six large groupings, three groupings of diseases, and three groupings of external causes (injuries): (i) CMNN, (ii) cancer, (iii) cardiovascular disease and other NCDs (CVD/NCDs), (iv) nonviolent injuries (road traffic accidents and other unintentional injuries), (v) suicides, and (vi) homicides. We also further divided these categories into 41 fine-grained groups. Table S1 contains details on the groupings and corresponding codes of ICD-10.

We addressed three critical challenges of vital registration data. First, we imputed missing age and sex (0.2% and 0.05% of deaths in Latin American cities and 0.005% and 0% of deaths in U.S. cities had missing information on age and sex, respectively) using a pro rata redistribution, based on cause of death, sex or age, country, and year. Second, we redistributed deaths assigned to ill-defined causes (4.1% for Latin American cities and 1.5% for U.S. cities) using a pro rata redistribution by age, sex, country, and year. Ill-defined causes are causes of death that do not provide useful information on the

cause of death for public health purposes, and we define these as ICD-10 codes R00-R94 and R96-R99. Third and last, we corrected for the underregistration of deaths at the city level by using an ensemble of death distribution methods. We assumed that U.S. cities had complete coverage of death counts and conducted this correction only for Latin American cities. More details on these three steps are available elsewhere (65). We also adjusted for the compositional effect of different age structures across cities, which have been shown to drive some scaling patterns (18) and which have a strong effect on mortality patterns. For this, we calculated the proportion of city residents aged 0 to 14, 15 to 39, 40 to 64, and 65 years and above, for the years of the study.

Statistical analysis

The basic scaling model (4) used for all analyses is a nonlinear power law relationship of the form

$$Y = Y_0 * N^\beta \quad (1)$$

We estimate these parameters by taking the natural logarithm of both sides of Eq. 1 and estimating an ordinary least squares (OLS) regression of the form

$$\ln(Y_{ij}) = \alpha + \beta * \ln(N_{ij}) + \epsilon_{ij} \quad (2)$$

where Y_{ij} is the number of deaths for the i th city in the j th country, α is a constant, β is the scaling exponent ($\beta < 1$ representing sublinear scaling, $\beta > 1$ superlinear scaling, and $\beta = 1$ linear scaling), N_{ij} is the population of the city, and ϵ is a residual (4). Data from different countries may present the same scaling pattern (same exponent β) but different magnitude (different α). Thus, to account for possible different levels of mortality rates and coding of causes of death by country, and to control for the role of age distribution in mortality, we also expanded the model in Eq. 2, based on previous work (33), as

$$\log(Y_{ij}) = \alpha + \beta * \log(N_{ij}) + \alpha_2 * \text{Country}_j + \alpha_3 * \text{Prop}(15_39)_{ij} + \alpha_4 * \text{Prop}(40_64)_{ij} + \alpha_5 * \text{Prop}(65p)_{ij} + \epsilon_{ij} \quad (3)$$

Here, the vector of variables Country_j refers to the country where the city is located, and $\text{Prop}(15_39)$, $\text{Prop}(40_64)$, and $\text{Prop}(65p)$ represent the percentage of the population in each city aged 15 to 39, 40 to 64, and 65+. We ran the model in Eq. 3 for all cities combined and also stratified these models by region to separately obtain scaling exponents (β) and intercepts (α) values for Latin American and U.S. cities separately. Last, we also leveraged the large number of cities in two Latin American countries to estimate country-specific scaling coefficients in Brazil ($n = 152$ cities), Mexico ($n = 92$), and all other Latin American cities ($n = 122$). Models for a single country (i.e., United States, Brazil, and Mexico) do not include any adjustment for country, while models for multiple countries (i.e., the models for all Latin American countries and for all Latin American countries minus Brazil and Mexico) include the adjustment covariates for country.

Last, to explore whether the general level and variability of each type of cause of death was associated with its scaling behavior, we followed the approach of Gomez-Lievano *et al.* (32). For this, we compared the scaling exponents (β) for each cause of death with the following: (i) their corresponding intercepts (α), which are a metric of the general levels of the phenomenon, and (ii) the square root of

the mean squared error, which are a metric of the variability of the phenomenon after accounting for city size, both obtained from Eq. 3 above. We explored the correlation between scaling patterns and general levels and variability by region, both graphically and by calculating Pearson's correlation coefficient.

We carried two sensitivity analyses. First, to test whether city definitions altered the scaling properties of causes of death, we explored two alternative definitions in the United States: (i) commuting zones (77) and (ii) an ad hoc definition designed to mimic the SALURBAL city definition (65, 71). This ad hoc definition was obtained by overlaying U.S. urban areas (block-level definition of urbanized areas) with U.S. counties and defining cities as agglomerations of counties that overlay with the city urbanized area. Details on these city definitions are available in table S6. Second, given the heterogeneity of city sizes across countries (78) and the potential differential importance of capital cities as economic hubs (60), we tested whether adding a covariate to Eq. 3 indicating whether the city was the largest in its country changed the inferences we observed in our main analysis. All analyses were conducted in R v4.1.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abl6325>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- UN-Habitat, Urbanization and development: Emerging futures; world cities report 2016, in *World Cities Report* (UN-Habitat, 2016).
- United Nations, *World Urbanization Prospects: The 2018 Revision* (United Nations, 2019); <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>.
- M. Franco, U. Bilal, A. V. Diez-Roux, Preventing non-communicable diseases through structural changes in urban environments. *J. Epidemiol. Community Health* **69**, 509–511 (2015).
- L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, G. B. West, Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7301–7306 (2007).
- L. M. A. Bettencourt, The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
- C. Dye, Health and urban living. *Science* **319**, 766–769 (2008).
- A. V. D. Roux, S. C. Slesinski, M. Alazraqui, W. T. Caiaffa, P. Frenz, R. J. Fuchs, J. J. Miranda, D. A. Rodriguez, O. L. Sarmiento Dueñas, J. Sirj, A. V. Vergara, A novel international partnership for actionable evidence on urban health in latin america: LAC-urban health and SALURBAL. *Global Chall.* **0**, 1800013 (2019).
- G. West, *Scale* (Orion Publishing Group, Limited, 2017).
- C. P. de Castro, G. F. dos Santos, A. D. de Freitas, M. I. dos Santos, R. F. S. Andrade, M. L. Barreto, Socio-economic urban scaling properties: Influence of regional geographic heterogeneities in Brazil. *PLOS ONE* **15**, e0242778 (2020).
- O. Patterson-Lomba, E. Goldstein, A. Gómez-Liévano, C. Castillo-Chavez, S. Towers, Per capita incidence of sexually transmitted infections increases systematically with urban population size: A cross-sectional study. *Sex. Transm. Infect.* **91**, 610–614 (2015).
- O. Patterson-Lomba, A. Gomez-Lievano, On the scaling patterns of infectious disease incidence in cities. arXiv:1809.00277 [physics.soc-ph] (2 September 2018).
- L. E. Rocha, A. E. Thorson, R. Lambiotte, The non-linear health consequences of living in larger cities. *J. Urban Health* **92**, 785–799 (2015).
- F. J. Antonio, S. de Picoli Jr., J. J. V. Teixeira, R. dos Santos Mendes, Growth patterns and scaling laws governing AIDS epidemic in Brazilian cities. *PLOS ONE* **9**, e111015 (2014).
- A. Gomez-Lievano, H. Youn, L. M. A. Bettencourt, The statistics of urban scaling and their connection to Zipf's law. *PLOS ONE* **7**, e40393 (2012).
- M. Oliveira, C. Bastos-Filho, R. Menezes, The scaling of crime concentration in cities. *PLOS ONE* **12**, e0183110 (2017).
- Y. S. Chang, W. J. Lee, J. H. Lee, Are there higher pedestrian fatalities in larger cities?: A scaling analysis of 115 to 161 largest cities in the United States. *Traffic Inj. Prev.* **17**, 720–728 (2016).
- S. B. Choi, Y. J. Lee, Y. S. Chang, Population size and urban health advantage: Scaling analyses of four major diseases for 417 US counties. *Int. J. Soc. Syst. Sci.* **10**, 35–55 (2018).
- J. Sutton, G. Shahtahmassebi, H. V. Ribeiro, Q. S. Hanley, Rural-urban scaling of age, mortality, crime and property reveals a loss of expected self-similar behaviour. *Sci. Rep.* **10**, 16863 (2020).

19. S. Tallman, Urban scaling and the impact of socioeconomic performance on health and environmental outcomes, thesis, University of New Mexico (2018).
20. J. Meirelles, C. R. Neto, F. F. Ferreira, F. L. Ribeiro, C. R. Binder, Evolution of urban scaling: Evidence from Brazil. *PLOS ONE* **13**, e0204574 (2018).
21. H. P. M. Melo, A. A. Moreira, É. Batista, H. A. Makse, J. S. Andrade, Statistical signs of social influence on suicides. *Sci. Rep.* **4**, 6239 (2014).
22. L. G. A. Alves, H. V. Ribeiro, E. K. Lenzi, R. S. Mendes, Distance to the scaling law: A useful approach for unveiling relationships between crime and urban metrics. *PLOS ONE* **8**, e69580 (2013).
23. C. A. Ignazzi, Scaling laws, economic growth, education and crime: Evidence from Brazil. *LSpace Geogr.* **43**, 324–337 (2014).
24. L. M. A. Bettencourt, J. Lobo, D. Strumsky, G. B. West, Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLOS ONE* **5**, e13541 (2010).
25. R. Louf, C. Roth, M. Barthelemy, Scaling in transportation networks. *PLOS ONE* **9**, e102007 (2014).
26. A. Loder, L. Ambühl, M. Menendez, K. W. Axhausen, Understanding traffic capacity of urban networks. *Sci. Rep.* **9**, 16283 (2019).
27. S. Sarkar, Urban scaling and the geographic concentration of inequalities by city size. *Environ. Plan. B* **46**, 1627–1644 (2018).
28. A. Alshamsi, E. Awad, M. Almhrezi, V. Babushkin, P. J. Chang, Z. Shoroye, A. P. Tóth, I. Rahwan, Misery loves company: Happiness and communication in the city. *EPJ Data Sci.* **4**, 7 (2015).
29. M. Ferguson, K. Ali, M. R. Olfert, M. Partridge, Voting with their feet: Jobs versus amenities. *Growth Chang.* **38**, 77–110 (2007).
30. T. Liu, J. Wang, Bringing city size in understanding the permanent settlement intention of rural–urban migrants in China. *Popul. Space Place* **26**, e2295 (2020).
31. D. R. Williams, C. Collins, Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Rep.* **116**, 404–416 (2001).
32. A. Gomez-Lievano, O. Patterson-Lomba, R. Hausmann, Explaining the prevalence, scaling and variance of urban phenomena. *Nat. Hum. Behav.* **1**, 0012 (2017).
33. M. Keuschnigg, S. Mutgan, P. Hedström, Urban scaling and the regional divide. *Sci. Adv.* **5**, eaav0042 (2019).
34. S. K. Chong, M. Bahrami, H. Chen, S. Balciyoy, B. Bozkaya, A. S. Pentland, Economic outcomes predicted by diversity in cities. *EPJ Data Sci.* **9**, 17 (2020).
35. M. Schläpfer, L. M. A. Bettencourt, S. Grauhan, M. Raschke, R. Claxton, Z. Smoreda, G. B. West, C. Ratti, The scaling of human interactions with city size. *J. R. Soc. Interface* **11**, 20130789 (2014).
36. C. Brelsford, J. Lobo, J. Hand, L. M. A. Bettencourt, Heterogeneity and scale of sustainable development in cities. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8963–8968 (2017).
37. D. B. de Castro, E. M. G. de Seixas Maciel, M. Sadahiro, R. C. Pinto, B. C. de Albuquerque, J. U. Braga, Tuberculosis incidence inequalities and its social determinants in Manaus from 2007 to 2016. *Int. J. Equity Health* **17**, 187 (2018).
38. P. H. Mullachery, D. A. Rodriguez, J. J. Miranda, N. López-Olmedo, K. Martinez-Folgar, M. L. Barreto, A. V. Diez Roux, U. Bilal, Mortality amenable to healthcare in Latin American cities: A cross-sectional study examining between-country variation in amenable mortality and the role of urban metrics. *Int. J. Epidemiol.* dyab137 (2021).
39. Centers for Disease Control and Prevention, Trends in tuberculosis-United States, 2011. *MMWR Morb. Mortal. Wkly Rep.* **61**, 181–185 (2012).
40. B. J. Baker, C. D. Jeffries, P. K. Moonan, Decline in tuberculosis among Mexico-born persons in the United States, 2000–2010. *Ann. Am. Thorac. Soc.* **11**, 480–488 (2014).
41. R. Stepler, M. H. Lopez, “U.S. Latino population growth and dispersion has slowed since onset of the great recession” (Pew Research Center, 2016).
42. T. J. Bollyky, T. Templin, M. Cohen, J. L. Dieleman, Lower-income countries that face the most rapid shift in noncommunicable disease burden are also the least prepared. *Health Aff.* **36**, 1866–1875 (2017).
43. J. J. Miranda, T. Barrientos-Gutiérrez, C. Corvalan, A. A. Hyder, M. Lazo-Porras, T. Oni, J. C. K. Wells, Understanding the rise of cardiometabolic diseases in low- and middle-income countries. *Nat. Med.* **25**, 1667–1679 (2019).
44. W. Mendoza, J. J. Miranda, Global shifts in cardiovascular disease, the epidemiologic transition, and other contributing factors: Toward a new practice of global health cardiology. *Cardiol. Clin.* **35**, 1–12 (2017).
45. K. A. Matthews, J. B. Croft, Y. Liu, H. Lu, D. Kanny, A. G. Wheaton, T. J. Cunningham, L. K. Khan, R. S. Caraballo, J. B. Holt, P. I. Eke, W. H. Giles, Health-related behaviors by urban-rural county classification—United States, 2013. *MMWR Surveill. Summ.* **66**, 1–8 (2017).
46. S. Rodríguez López, U. Bilal, A. F. Ortigoza, A. V. Diez-Roux, Educational inequalities, urbanicity and levels of non-communicable diseases risk factors: Evaluating trends in Argentina (2005–2013). *BMC Public Health* **21**, 1572 (2021).
47. A. Ruiz-Alejós, R. M. Carrillo-Larco, J. J. Miranda, C. A. M. Anderson, R. H. Gilman, L. Smeeth, A. Bernabé-Ortiz, Addressing the impact of urban exposure on the incidence of type 2 diabetes mellitus: The PERU MIGRANT Study. *Sci. Rep.* **8**, 5512 (2018).
48. R. M. Carrillo-Larco, A. Bernabé-Ortiz, T. D. Pillay, R. H. Gilman, J. F. Sanchez, J. A. Poterico, R. Quispe, L. Smeeth, J. J. Miranda, Obesity risk in rural, urban and rural-to-urban migrants: Prospective results of the PERU MIGRANT study. *Int. J. Obes.* **40**, 181–185 (2016).
49. A. Taype-Rondan, A. Bernabé-Ortiz, G. F. Alvarado, R. H. Gilman, L. Smeeth, J. J. Miranda, Smoking and heavy drinking patterns in rural, urban and rural-to-urban migrants: The PERU MIGRANT Study. *BMC Public Health* **17**, 165 (2017).
50. V. Akuraju, P. Pradhan, D. Haase, J. P. Kropp, D. Rybski, Relating SDG11 indicators and urban scaling—An exploratory study. *Sustain. Cities Soc.* **52**, 101853 (2020).
51. M. Fragkias, J. Lobo, D. Strumsky, K. C. Seto, Does size matter? Scaling of CO2 emissions and U.S. Urban areas. *PLOS ONE* **8**, e64727 (2013).
52. L. Han, W. Zhou, S. T. Pickett, W. Li, L. Li, An optimum city size? The scaling relationship for urban population and fine particulate (PM2.5) concentration. *Environ. Pollut.* **208**, 96–101 (2016).
53. N. Gouveia, J. L. Kephart, I. Dronova, L. McClure, J. T. Granados, R. M. Betancourt, A. C. O’Ryan, J. L. Texcalac-Sangrador, K. Martinez-Folgar, L. Rodriguez, A. V. Diez-Roux, Ambient fine particulate matter in Latin American cities: Levels, population exposure, and associated urban factors. *Sci. Total Environ.* **772**, 145035 (2021).
54. B. D. Fulton, Health care market concentration trends in the united states: Evidence and policy responses. *Health Aff.* **36**, 1530–1538 (2017).
55. M. E. Cyr, A. G. Etchin, B. J. Guthrie, J. C. Benneyan, Access to specialty healthcare in urban versus rural US populations: A systematic literature review. *BMC Health Serv. Res.* **19**, 974 (2019).
56. J. J. Salinas, S. Al Snih, K. Markides, L. A. Ray, R. J. Angel, The rural-urban divide: Health services utilization among older Mexicans in Mexico. *J. Rural Health* **26**, 333–341 (2010).
57. H. Hong, O. J. Mújica, J. Anaya, V. C. Lansingh, E. López, J. C. Silva, The challenge of universal eye health in latin America: Distributive inequality of ophthalmologists in 14 countries. *BMJ Open* **6**, e012819 (2016).
58. M. Barraza-Lloréns, S. Bertozzi, E. González-Pier, J. P. Gutiérrez, Addressing inequity in health and health care in Mexico. *Health Aff.* **21**, 47–56 (2002).
59. N. Afshar, D. R. English, R. L. Milne, Rural–urban residence and cancer survival in high-income countries: A systematic review. *Cancer* **125**, 2172–2184 (2019).
60. E. Arcaute, E. Hatna, P. Ferguson, H. Youn, A. Johansson, M. Batty, Constructing cities, deconstructing scaling laws. *J. R. Soc. Interface* **12**, 20140745 (2015).
61. D. Pumain, in *Complexity Theories of Cities Have Come of Age* (Springer, 2012), pp. 91–103.
62. R. Jedwab, D. Vollrath, *The Mortality Transition, Malthusian Dynamics, and the Rise of Poor Mega-Cities* (The George Washington University, Institute for International Economic Policy, 2015).
63. A. R. Omran, The epidemiologic transition: A theory of the epidemiology of population change. 1971. *Milbank Q.* **83**, 731–757 (2005).
64. S. M. Barreto, J. J. Miranda, J. P. Figueroa, M. I. Schmidt, S. Munoz, P. P. Kuri-Morales, J. B. Silva, Epidemiology in Latin America and the Caribbean: Current situation and challenges. *Int. J. Epidemiol.* **41**, 557–571 (2012).
65. U. Bilal, P. Hessel, C. Perez-Ferrer, Y. L. Michael, T. Alfaro, J. Tenorio-Mucha, A. A. L. Friche, M. F. Pina, A. Vives, H. Quick, M. Alazraqi, D. A. Rodriguez, J. J. Miranda, A. V. Diez-Roux; SALURBAL group, Life expectancy and mortality in 363 cities of Latin America. *Nat. Med.* **27**, 463–470 (2021).
66. C. J. L. Murray, S. C. Kulkarni, C. Michaud, N. Tomijima, M. T. Bulzacchelli, T. J. Landiorio, M. Ezzati, Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLOS Med.* **3**, e260 (2006).
67. J. Lobo, L. M. Bettencourt, D. Strumsky, G. B. West, Urban scaling and the production function for cities. *PLOS ONE* **8**, e58407 (2013).
68. R. L. Moomaw, Productivity and city size: A critique of the evidence. *Q. J. Econ.* **96**, 675–688 (1981).
69. C. Sarkar, C. Webster, Healthy cities of tomorrow: The case for large scale built environment–health studies. *J. Urban Health* **94**, 4–19 (2017).
70. A. Peralta, J. Benach, C. Borrell, V. Espinel-Flores, L. Cash-Gibson, B. L. Queiroz, M. Mari-Dell’Olmo, Evaluation of the mortality registry in Ecuador (2001–2013) – Social and geographical inequalities in completeness and quality. *Popul. Health Metrics* **17**, 3 (2019).
71. D. A. I. Quistberg, A. V. D. Roux, U. Bilal, K. Moore, A. Ortigoza, D. A. Rodriguez, O. L. Sarmiento, P. Frenz, A. A. Friche, W. T. Caiaffa, A. Vives, J. J. Miranda; SALURBAL Group, Building A data platform for cross-country urban health studies: The SALURBAL study. *J. Urban Health* **96**, 311–337 (2019).
72. R. E. McKeown, The epidemiologic transition: Changing patterns of mortality and population dynamics. *Am. J. Lifestyle Med.* **3**, 195–265 (2009).
73. L. M. A. Bettencourt, V. C. Yang, J. Lobo, C. P. Kempes, D. Rybski, M. J. Hamilton, The interpretation of urban scaling analysis in time. *J. R. Soc. Interface* **17**, 20190846 (2020).
74. J. Depersin, M. Barthelemy, From global scaling to the dynamics of individual cities. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2317–2322 (2018).

75. National Center for Health Statistics, Compressed Mortality File, 2012–2016, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program, (2018).
76. World Health Organization, WHO methods and data sources for country-level causes of death 2000–2019 (Global Health Estimates Technical Paper WHO/DDI/DNA/GHE/2020.2; WHO, Geneva, 2020); https://cdn.who.int/media/docs/default-source/gho-documents/global-health-estimates/ghe2019_cod_methods.pdf.
77. C. S. Fowler, L. Jensen, Bridging the gap between geographic concept and the data we have: The case of labor markets in the USA. *Environ. Plan. A* **52**, 1395–1414 (2020).
78. S. Arshad, S. Hu, B. N. Ashraf, Zipf's law and city size distribution: A survey of the literature and future research agenda. *Phys. A* **492**, 75–92 (2018).

Acknowledgments: We acknowledge the contribution of all coinvestigators of the SALURBAL project. For more information on SALURBAL and to see a full list of investigators, see <https://drexel.edu/lac/salurbal/team/>. SALURBAL acknowledges the contributions of many different agencies in generating, processing, facilitating access to data, or assisting with other aspects of the project. Please visit <https://drexel.edu/lac/data-evidence> for a complete list of data sources. The findings of this study and their interpretation are the responsibility of the authors and do not represent the views or interpretations of the institutions or groups that compiled, collected, or provided the data. The use of data from these institutions does not claim or imply that they have participated in, approved, endorsed, or otherwise supported the development of this publication. They are not liable for any errors, omissions, or other defect or for any actions taken in reliance thereon. **Funding:** This project was supported by the Office of the Director of the NIH under award number DP5OD26429 and by

the Wellcome Trust initiative, “Our Planet, Our Health” (grant 205177/Z/16/Z). The funding agencies had no involvement in the study design; in the data collection, analyses, or interpretation of data; in the writing of this work; or in the decision to submit the manuscript for publication. **Author contributions:** A.V.D.R., D.A.R., J.J.M., M.L.B., and U.B. obtained funding for this study. U.B., A.V.D.R., C.P.d.C., G.F.d.S., and R.F.S.A. conceived the study. U.B. conducted the statistical analysis and wrote the first draft of the manuscript with support from A.V.D.R., C.P.d.C., and R.F.S.A. T.A., T.B.-G., K.M.-F., and J.J.M. contributed to data collection. All authors interpreted results, reviewed the manuscript for critical content, and approved the final version. U.B. is the guarantor for this study. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Vital registration data for Brazil, Chile, Colombia, and Mexico were downloaded from publicly available repositories from statistical agencies in each country. Vital registration data for Argentina, Costa Rica, El Salvador, Panama, and Peru were obtained directly from statistical agencies in each country. For a link to these agencies, websites can be accessed via <https://drexel.edu/lac/data-evidence/data-acknowledgements/>. Data on vital registration for the United States were obtained under request from the National Center for Health Statistics, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Code and data used in this study are available at <https://zenodo.org/record/5553365>. The public-use dataset has had death counts suppressed if the number was below 10, due to data sharing limitations.

Submitted 26 July 2021

Accepted 20 October 2021

Published 8 December 2021

10.1126/sciadv.abl6325