Growth scaling for the early dynamics of HIV/AIDS epidemics in Brazil and the influence of socio-demographic factors

L. Dinh\textsuperscript{a,}\textsuperscript{*}, G. Chowell\textsuperscript{a,b}, R. Rothenberg\textsuperscript{a}

\textsuperscript{a}Division of Epidemiology and Biostatistics, School of Public Health, Georgia State University, Atlanta, GA, USA
\textsuperscript{b}Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

\section*{A R T I C L E   I N F O}

Article history:
Received 26 July 2017
Revised 25 December 2017
Accepted 29 December 2017
Available online 9 January 2018

Keywords:
Sub-exponential growth
HIV/AIDS epidemic
Demographic factors
Brazil

\section*{A B S T R A C T}

The early dynamics of an infectious disease outbreak can be affected by various factors including the transmission mode of the disease and host-specific factors. While recent works have highlighted the presence of sub-exponential growth patterns during the early phase of epidemics, empirical studies examining the contribution of different factors to early epidemic growth dynamics are lacking. Here we aim to characterize and explain the early incidence growth patterns of local HIV/AIDS epidemics in Brazil as a function of socio-demographic factors. For this purpose, we accessed annual AIDS incidence series and state-level socio-demographic variables from publicly available databases. To characterize the early growth dynamics of the HIV/AIDS epidemic, we employed the generalized-growth model to estimate with quantified uncertainty the scaling of growth parameter (p) which captures growth patterns ranging from constant incidence (p = 0) to sub-exponential (0 < p < 1) and exponential growth dynamics (p = 1) at three spatial scales: national, regional, and state levels. We evaluated the relationship between socio-demographic variables and epidemic growth patterns across 27 Brazilian states using mixed-effect regression analyses. We found wide variation in the early dynamics of the AIDS epidemic in Brazil, displaying sub-exponential growth patterns with the p parameter estimated substantially below 1.0. The mean p was estimated to be 0.81 at the national level, with a range of 0.72–0.85 at the regional level, and a range of 0.28–0.96 at the state level. Our findings support the notion that socio-demographic factors contribute to shaping the early growth dynamics of the epidemic at the local level. Gini index and socio-demographic index were negatively associated with the parameter p, whereas urbanicity was positively associated with p. The results could have theoretical significance in understanding differences in growth scaling across different sexually transmitted disease systems, and have public health implications to guide control.

\section*{1. Introduction}

Infectious diseases represent one of the most important threats to humans. Thus, a better understanding of the factors that drive the epidemic growth dynamics during the first few generations of disease transmission is needed to construct more accurate mechanistic models, and in turn, generate improved disease forecasts. Indeed, the importance of such factors is expected to differ across disease systems (e.g., from influenza, Ebola, to HIV/AIDS). Previous research has shown that the early growth profile of epidemic outbreaks follows sub-exponential growth dynamics with varying polynomial degrees across historic and contemporary outbreaks involving different infectious diseases (Colgate et al., 1989; Chowell et al., 2015; Chowell et al., 2016; Chowell et al., 2016; Valeri et al., 2016; Viboud et al., 2016). For instance, the cumulative number of HIV/AIDS cases in the United States in the 1980s grew as a cubic polynomial (sub-exponential) in time rather than exponentially (May and Anderson 1987; Colgate et al., 1989), and more recently the Ebola epidemic in administrative areas of West Africa displayed sub-exponential growth dynamics (Chowell et al., 2015).

Epidemic models that largely assume homogenous and well-mixed populations support exponential growth dynamics in the absence of behavior change or control interventions (Chowell et al., 2016). On the other hand, sub-exponential epidemic growth can arise from a combination of mechanisms including: (1) spatial constraints (e.g., clustering) associated with the transmission mode of the disease, (2) population heterogeneities affecting susceptibility and infectivity at the individual level, and (3) reactive behavior changes and early onset of interventions that gradually mitigate the transmission rate (Szendroi and Csanyi 2004; Chowell et al., 2015). To characterize epidemic growth profiles, the generalized-
growth model is useful to quantify the “scaling of growth” relative to the exponential growth pattern that is expected with a constant growth rate (Viboud et al., 2016). Recent research examining the early ascending phase of infectious disease outbreaks evidenced early sub-exponential epidemic growth dynamics, ranging from very slow (e.g. for the 2014 Ebola outbreak in Bomi, Liberia) to nearly exponential growth (e.g. the 1972 smallpox outbreak in Khulna and the 1918 influenza pandemic in San Francisco) (Chowell et al., 2016; Viboud et al., 2016). However, empirical studies characterizing the scaling of epidemic growth at different spatial scales and analyzing the association between specific growth profiles and socio-demographic variables are lacking.

Here we sought to examine the variation in epidemic growth profiles of the HIV/AIDS epidemic at three spatial scales in Brazil, a country with diverse patterns of urbanization, population density, socio-demographic development, and income inequality (Incidi et al., 2000). For this purpose, we quantified the relationship between early incidence growth dynamics and socio-demographic indicators as well as the timing of epidemic onset. The HIV/AIDS epidemic in Brazil is spatially heterogeneous (Antonio et al., 2014), and hence, it offers a unique opportunity to investigate incidence growth profiles at different spatial scales using epidemiological data collected through comprehensive epidemiological surveillance (The Brazilian Ministry of Health; UNAIDS, 2004). In Brazil, the first AIDS case was identified in 1982, with affected patients initially subject to intense stigma and discrimination. However, Brazil’s political transformation, known as “abertura” (1974–1985), from an authoritarian military dictatorship to a democratic government (Celentano and Beyrer, 2008), together with extensive social movements, significantly improved its response to the AIDS epidemic (Bielh 2007). Brazil established the National AIDS Program in 1986 and a universal provision of antiretroviral drugs (ARVs) in 1996.

2. Materials and methods

2.1. Data sources

2.1.1. Epidemiological data

To characterize the early incidence growth patterns of the AIDS epidemic in Brazil, we accessed annual AIDS incidence series during 1982–1999 at three different spatial scales (national, regional and state) from the DATASUS database of Department of Informatics of the Brazilian Health System (Department of Informatics of the Brazilian Health System). The AIDS national surveillance system in Brazil officially started in 1987 (Ramos et al., 2011), relying on mandatory AIDS case reporting (Oliver Bacon, Maria Lucia Pecoraro et al. 2014). Each Brazilian state recorded cases occurring before 1986 retrospectively without following a specific standard (Ramos et al., 2011). For surveillance purposes, Brazil initially adopted the CDC AIDS case definition of 1985 (Ramos et al., 2011), but later implemented its own case definitions (Campos et al., 2005).

The incubation period from HIV infection to the development of AIDS varies depending on intrinsic characteristics of patients, especially age at which infection occurs (Bacchetti and Moss 1989; Rosenberg et al., 1994). For adults, the incubation period has been estimated to have a median of approximately 10 years in a non-parametric model (Bacchetti and Moss 1989), and a mean of approximately 13 years in a model using a scaled gamma distribution (Bailey 1997). Nevertheless, when using AIDS incidence in lieu of HIV data, we were based on the assumption that the “diagnostic lag” of HIV cases, which consists of the incubation period and the diagnostic lag after developing AIDS, averaged out. For each incidence curve, the onset time corresponds to the year of the first HIV/AIDS case report. At the national level, the first case was reported in 1982.

2.1.2. Socio-demographic data

We obtained the following socio-demographic variables for 27 Brazilian states: population density, urbanicity, Gini index, and the socio-demographic index from the 1991 National Census data, which is administered by the Brazilian Institute of Geography and Statistics (IBGE) (Brazilian Institute of Geography and Statistics (IBGE)) and the Global Burden of Disease Study, led by the Institute for Health Metrics and Evaluation (IHME) (Institute for Health Metrics and Evaluation (IHME)). Socio-demographic variables, descriptive statistics, and their sources are summarized in Table 1 while their corresponding distributions are shown in Fig. 1.

2.2. Characterizing the early epidemic growth phase

Slower-than-exponential epidemic growth is expected for HIV/AIDS spread as transmission is driven by direct contact (Colgate et al., 1989). In order to capture the possibility of sub-exponential growth dynamics, we employed the generalized growth model (GGM) (Chowell and Viboud 2016; Viboud et al., 2017), which describes incidence growth using a single differential equation:

\[ C'(t) = rC(t)^p \]

where \( C'(t) \) denotes the incidence at time \( t \), and the solution \( C(t) \) denotes the cumulative number of cases at time \( t \). Parameter \( r \) describes the growth rate, and \( p \in [0, 1] \) is the deceleration of growth parameter. If \( p = 0 \), the equation describes an epidemic with constant incidence, while \( p = 1 \) models exponential growth dynamics, as described in refs (Chowell et al., 2016; Viboud et al., 2016). It is worth-noting that the growth rate \( r \) is a so-called inno-cent parameter (Nasell 2003) as it could be eliminated by rescaling time. Here we focus on the variation of the scaling of growth parameter \( p \), which is the only essential parameter in the GGM.

2.2.1. Parameter estimation and confidence intervals

We calibrated the generalized growth model to the early epidemic growth phase of incidence case data comprising the first few disease generations as in previous studies (Nishiura 2010; Viboud et al., 2016). Specifically, the initial epidemic growth phase comprising 10 epidemic years was selected for our analyses. It is important to note that national access to antiviral drugs (ARVs) in Brazil started in 1996, which has further delayed the progression of the epidemic (Teixeira 2002; Gougeon 2003). Second, the time lag from diagnosis to reporting for the AIDS epidemic in Brazil has been estimated at 2 years (Barbosa and Struchiner 2002; Costa and Luiz 2003). As such, our study period mostly reflects the transmission dynamics of the epidemic prior to implementation of ARV treatment in Brazil.

Model parameters \( p \) and \( r \) were estimated jointly by non-linear least squares curve fitting using Trust-reflective-region Least Squares algorithm implemented in MATLAB (The Mathworks, Inc.). The initial parameter \( C(0) \) was fixed according to case incidence at the onset year for the corresponding spatial scale. Parameter uncertainty was quantified by simulating 200 realizations of best-fit curve \( C(t) \) using parametric bootstrap with a Poisson error structure, as in a previous study (Viboud et al., 2016). Confidence intervals for parameters \( p \) and \( r \) were then computed by taking 2.5th and 97.5th percentile points.

2.2.2. Models comparison

We assessed the goodness of fit provided by the generalized-growth model and the simpler exponential growth model using the
Table 1
Summary statistics of the socio-demographic predictor variables and their sources.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Description</th>
<th>Descriptive Statistics</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centered time of onset</td>
<td>year</td>
<td>Time of onset minus 1980</td>
<td>Mean: 4.41, SD: 2.71</td>
<td>DATASUS (15)</td>
</tr>
<tr>
<td>Population density</td>
<td>pop_den</td>
<td>Number of people per kilometer square</td>
<td>Mean: 50.41, SD: 74.38</td>
<td>Microdata of the Demographic Census (IBGE) (16)</td>
</tr>
<tr>
<td>Urbanicity</td>
<td>urbanicity</td>
<td>Proportion (in decimal fraction) of people living in the urban area of the state</td>
<td>Mean: 0.70, SD: 0.13</td>
<td></td>
</tr>
<tr>
<td>Gini index</td>
<td>gini</td>
<td>A measure of inequality, obtained from the per capita household income, Range from 0 - complete equality - to 1 - only one person concentrates the income.</td>
<td>Mean: 0.62, SD: 0.02</td>
<td></td>
</tr>
<tr>
<td>Socio demographic index</td>
<td>sdi</td>
<td>A measure of socio-demographic development, based on average income per person, educational attainment, and total fertility rate (TFR). Range from 0 - the lowest income per capita, lowest educational attainment, and highest TFR - to 1 - the highest income per capita, highest educational attainment, and lowest TFR (17)</td>
<td>Mean: 0.48, SD: 0.08</td>
<td>Global Burden of Disease Study 2015 (GDB 2015) (17)</td>
</tr>
</tbody>
</table>

Fig 1. The distribution of socio-demographic predictors across the Brazilian states in 1991. The level of predictors is indicated by the ascending color intensity. The range of each predictor is specified in the corresponding color legend bar.

The residual standard error, which is given by:

$$RSE = \sqrt{ \frac{RSS}{n - k - 1}}$$

where RSS is the residuals sums squared obtained from the best fit curve and observed case series data, n – k – 1 are the degrees of freedom of a model with n data points and k estimated parameters to penalize for model complexity (Valeri et al., 2016). For comparison, we computed the ratio of RSEs (hereafter denoted by RRSE) obtained from the generalized growth model and the exponential growth model. RRSE < 1 indicates that the generalized growth model enhanced model fitting to the data, and vice versa.

2.3. Early growth and time of onset of the AIDS epidemic across Brazilian states

To examine the association between the time of onset and the deceleration parameter of growth $p$, we visualized the parameter $p$ against time of onset. Centered time of onset was spatially visualized to explore possible clustering patterns in the spread of the AIDS epidemic across Brazilian states. We tested for spatial auto-
correlation of time of onset, using Moran’s I test. Contiguity was defined as areas sharing any boundary point. Moran’s I was computed using 999 Monte-Carlo simulations. Analysis and visualization were conducted with R v3.3.2

2.4. Quantifying the association between socio-demographic variables and the deceleration of growth parameter \( p \) across Brazilian states

To analyze the association between socio-demographic variables and the variability in the deceleration parameter \( p \) at the state level, we employed a mixed-effect regression model (Austin et al., 2001). The predictors in the model were centered time of onset, Gini index, population density, urbanicity, and the social demographic index. Centered time of onset was included in the model simply as the controlling factor for the difference of the Census time (1991) relative to AIDS epidemic period used to calibrate the parameter \( p \).

We considered the full uncertainty in our \( p \) estimates and included all \( p \) estimates derived from bootstrapped simulations (200 for each state) at the individual level.

\[
p_{ij} = \beta_0 + \beta X_j + \xi_j + e_{ij}^s, \quad e_{ij}^s \sim N(0, \sigma^2)
\]

where \( p_{ij} \) denotes the obtained \( p \) for simulation \( s \) with \( s = 1, 2, \ldots, 200 \) for state \( j \) \((j = 1, 2, \ldots, 27)\); \( \beta_0 \) is the true parameter \( p \) for state \( j \); \( e_{ij} \) is the error term in simulation \( s \) for state \( j \).

The predictors were entered in the regression model at the cluster level, taking the following form:

\[
p_j = \beta_0 + \beta X_j + \xi_j \sim N(0, \sigma^2)
\]

where \( X_j \) denotes the group of selected predictors; \( \beta_0 \) is intercept and \( \beta \) is the vector of parameters that express the relation between predictors and \( p_j \); \( \xi_j \) is error term in the regression equation between predictors and \( p_j \) \((j = 1, 2, \ldots, 27)\).

By combining the Eqs. (1) and (2), we obtained:

\[
p_{ij} = \beta_0 + \beta X_j + \xi_j + e_{ij}^s
\]

Eq. (3) provides the final form of the regression model that incorporates random effects by treating the state-specific intercept as a random variable.

As the deceleration parameter \( p \) is a continuous variable in the interval \([0,1]\), it requires an appropriate transformation or distribution. The common distributions for a model response variable of \((0,1)\) are logit transformation and beta regression, but others are available. We fitted the parameter \( p \) to a number of distributions using the gamlss package in R, and selected the logistic distribution based on the Akaika Information Criteria. Then, we created a truncated distribution of logistic distribution, using gen.trun command (Rigby and Stasinopoulos 2009). The mixed effect regression model (Eq. (3)) was fitted to data using galmss command, with random() option and the truncated distribution. All analyses were conducted with galmss package, on R v3.3.2.

3. Results

Our analysis of the early incidence growth phase of the AIDS epidemic in Brazil revealed a diversity of sub-exponential growth profiles at the national, regional and state levels, as shown in Fig. 3. Importantly, based on RSE criteria, we found that the generalized-growth model yielded significantly enhanced fits to the AIDS incidence data compared to those derived using the exponential growth model (RRSE: 0.12–0.75), except for Amapá (RRSE = 1.07). Our results from the regression analysis indicated that population density, socio-demographic index, and Gini index were negatively associated with the “deceleration of growth” parameter \( p \), while urbanicity was positively associated with \( p \).

3.1. Early incidence growth patterns at different spatial scales

Representative fits of the generalized growth model to national and regional HIV/AIDS epidemics in Brazil are displayed in Fig. 4. At the national level, the parameter \( p \) was estimated at 0.81 (95%CI: 0.80–0.82). As shown, North, Northeast, and Midwest regions shared a pattern of intermediate growth profiles \((p = 0.7)\) with a slight difference in uncertainty as shown in Fig. 4 and Supplement. For the Southeast and South regions, high values of \( p \) above 0.8 were estimated \((p = 0.81 \text{ and } 0.85 \text{ respectively})\), which are in line with our national estimates.

Importantly, the deceleration of growth parameter \( p \) displayed substantial variation across Brazilian states and wider uncertainty in slowly growing epidemics (Supplementary 2). Estimates of \( p \) varied from 0.28 (95% CI: 0.14–0.43) for Acre, indicating a slow epidemic growth pattern, to 0.96 (95% CI: 0.82, 1) for Amapá, consistent with near-exponential growth dynamics. These two states located in northern Brazil exhibited the largest difference in mean \( p \) estimates. Slow-growth profiles with \( p \) smaller than 0.5 were also detected in Roraima and Alagoas \((p = 0.38\) and 0.43, respectively). There were high values of \( p \) above 0.80 in Rondônia, Ceará, Mato Grosso do Sul, São Paulo and Paraná. The growth pattern of other states, in general, correlates well with that of the region they belong to, displaying growth profiles ranging from moderate to intermediate sub-exponential growths.

3.2. Early growth and time of onset of the AIDS epidemic across Brazilian states

In the scatterplot displayed in Fig. 5, the states that started earlier in the epidemic had a much more focused \( p \), between 0.6 and 0.8, with some even higher, but none below 0.6 from 1981 to 1986. By contrast, the value of \( p \) showed significantly more variation among states with year of onset after 1986. Moreover, as can be seen in Fig. 2, the timing of onset of the AIDS epidemic at the state level in Brazil clearly displays spatial clusters (Moran’s I statistic = 0.20, \( p \text{-value} = 0.04 \)). The Southern cluster, with São Paulo at the center, tended to report AIDS cases earlier than northern cluster. At the same time, as reported in the previous section, the AIDS epidemic in southern regions had higher \( p \) than in northern ones at the regional level.
3.3. The association between socio-demographic variables and the deceleration parameter $p$ across Brazilian states

Results from sequentially fitting the mixed-effect regression models are presented in Table 2. The most noticeable finding is that the association between sociodemographic index and the parameter $p$ changed from positive to negative after Gini index was included in the model. Specifically, in the simple model with only ‘year’ and SDI (Model 2) and the model adding population density (Model 3), coefficient estimates of SDI were 0.330 and 0.968, respectively. In contrast, when Gini index (Model 4) and urbanicity (Model 5) were sequentially added to the model, coefficient estimates of SDI changed to negative, being $-0.427$ and $-1.977$, respectively. Moreover, the coefficient of population density was pulled towards zero when controlling for socioeconomic and Gini indices albeit the sign of coefficient remained negative.

Table 3 displays results of the final regression model, with all predictors of interest included. Our regression analysis found a statistically significantly positive association between the proportion of urbanicity and the deceleration parameter $p$ (coefficient estimate (est.) = 1.116, $p < 0.0001$). Meanwhile, population density, sociodemographic index and Gini index were negatively associated with $p$ (est$_{pop\_den}$ = $-0.00038$, est$_{pop\_den} < 0.0001$, est$_{gini}$ = $-6.753$, est$_{gini} < 0.0001$; est$_{sdi}$ = $-1.977$, est$_{sdi} < 0.0001$).

4. Discussion

Our results reveal spatial heterogeneity of growth patterns of the AIDS epidemic in Brazil, and suggest the crucial role of socio-demographic variables in explaining that variation. We found a substantial variation of the deceleration of growth parameter “$p$” across geographical areas, ranging from very slow to nearly-exponential growth. Our analysis at the state level revealed that urbanicity was positively associated with $p$ while Gini index and socio-demographic index were negatively associated with this parameter.
We observed sub-exponential growth dynamics during the early phase of this epidemic. This finding further supports the idea of sub-exponential growth of AIDS epidemics in the earlier works (May and Anderson 1987; Colgate et al., 1989; Viboud et al., 2016). Unlike Japan and New York City where the early growth patterns were approximately linear (p ~ 0.5) (Chowell et al., 2016), the scaling of growth for the epidemic in Brazil was somewhat higher (p = 0.81, 95%CI: 0.80–0.82) at the national level. This is in agreement with the fact that the AIDS epidemic in Brazil ranked second in AIDS burden worldwide (Inciardi et al., 2000) particularly during our study period.

At the regional level, we found relatively wide variation in epidemic growth patterns. We estimated intermediate scaling of growth in the North, Northeast and Midwest (p ~ 0.7), as opposed to the higher scaling of growth estimated for the South and Southeast (p ~ 0.85 and 0.81, respectively). This disparity may be explained by the economic, social, cultural and behavioral diversity of Brazil – as it is often said “there are several Brazils” (Antunes 2001). In fact, Brazil is an enormous country consisting of small indigent villages in some northerly states and cosmopolitan cities in the southeast (Inciardi et al., 2000; Antunes 2001).

Epidemic incidence patterns encompassing large spatial scales (e.g., national) can mask substantial heterogeneities that are only evident at smaller spatial scales (e.g., county or administrative area). During the 2014–16 Ebola epidemic, national incidence curves for Guinea, Sierra Leone, and Liberia displayed short-lived exponential growth periods. Yet, local epidemics at the level of administrative areas were asynchronous and largely followed polynomial growth dynamics during 3–4 generations of disease transmission (Chowell et al., 2015). In this study, epidemics at the national and regional levels result from the aggregation of smaller sub-epidemics at the state level. Indeed, our GGM model provided a good fit to those subnational epidemics (i.e. 95% confidence intervals capture the real data) (Supplement Fig 1) but displayed substantial variation in growth patterns across states. This subnational variation in growth patterns often negatively affected goodness of fit at the national level and regional levels (see e.g., panels for Brazil, North, and Southeast regions in Fig. 4).

Variability in epidemic growth across states ranged from very slow sub-exponential growth epidemics to nearly exponential ones. Especially, the largest discrepancy across states was observed within the North – between Amapá and Acre. The nearly exponential growth in Amapá may be due to its special geographical and economic characteristics (Parriault et al., 2015). As Amapá borders with French Guinea, which has much higher GPD and is a supply and transit hub for illegal gold mining, sex work has been widely present in this area (Parriault et al., 2015), which has likely facilitated the spread of HIV/AIDS.

We also observed an association between the time of onset of the AIDS epidemic at the state level and the parameter p. The states where the AIDS epidemic started early generally had higher and more focused values of p in the range 0.6–0.9, while those states with later onsets showed higher variation in the p parameter. This suggests that most of the heterogeneity in p is associated with the state-level epidemics with later onset (1987–1991) (Fig 5). It is important to note that epidemic growth patterns for any infectious disease system can be affected by reporting delays. In our study, epidemic growth patterns and the epidemic onset for sub-epidemics starting before 1986 may have been affected by either underreporting or reporting delays. In our regression analyses, by including the “year” variable in our regression model, we examined the association between socio-demographic factors and the deceleration parameter p after controlling for the effect of the epidemic onset time, and for the time of the Census associated with our socio-demographic variables.

Epidemic growth patterns are also shaped by intrinsic factors relating to the natural history of the disease (e.g. transmission mode, the variability of the incubation period) and the particular characteristics of the geographic setting where the epidemic takes place. The main drivers of the HIV/AIDS epidemic in Brazilian urban areas have been linked to a high prevalence of drug users, sex workers and female transgender persons (Ellison et al., 1993; Inciardi et al., 2000; UNAIDS, 2004; Malta et al., 2008). Indeed, we found a positive correlation between urbanicity and the parameter p whereby higher urbanicity in the south of Brazil is correlated with higher estimates of the scaling of growth parameter p. A possible explanation for this pattern is out-migration, which is a negative feature of urbanization. Specifically, as a city evolves, low-income residents may leave areas of gentrification or concentrate in a decaying city core (Vlahov and Galea 2002). For instance, the concentration of an economically deprived population in rundown areas can form a spatial contact structure that facilitates risky behaviors (e.g. needle-sharing, sex trade, multiplex relationships) of HIV/AIDS transmission (Inciardi et al., 2000).

We also found an inverse association between SDI and the deceleration parameter p. This is consistent with the conventional

---

**Fig 4.** Estimates and 95% confidence intervals for parameters r and p obtained by nonlinear least-square fitting the generalized growth model as explained in the text, at national and regional levels. The model was fitted to the initial epidemic growth phase comprising the first 10 years of AIDS case reports. Case incidence (y-axis) is presented in log scale to zoom in changes in the early phase.
belief that a higher level of socio-demographic development may curb the growth of the HIV/AIDS epidemic. For instance, people in more socio-demographically developed areas are more likely to readily adopt behavior changes and benefit from control and prevention interventions (Springborn et al., 2015). When preventative measures are adopted differentially and gradually within and between states, heterogeneity in their susceptibility and infectivity can be amplified. Thus, changes in environmental factors can lead to population heterogeneities that can give rise to sub-exponential epidemic growth profiles.

Regarding Gini index (a measure of income inequality) and population density, we found these predictors to be negatively related to the deceleration parameter $p$, once other covariates were controlled for. It is important to note that previous empirical studies have reported higher HIV/AIDS prevalence in areas of higher population density and higher Gini index (Holmqvist 2009; Brodish 2015). However, the deceleration parameter is a measurement of dynamic growth processes rather than a snapshot of a condition as “prevalence” is. The negative association between Gini index and $p$ can be explained by the fact that the Gini coefficient captures evenness and spatial segregation of different subpopulations in an area (Iceland et al., 2002). As such, areas associated with higher Gini index may be associated with higher clustering levels, imposing “spatial constraints” on HIV/AIDS transmission dynamics. Hence, we hypothesize that disease spreads more slowly in areas with higher Gini indices.

As for population density, although the association with $p$ was statistically significant ($p < 0.0001$), its magnitude was small (est. $= -0.00038$). For an increase of 1 person/square kilometer in population density, there was a corresponding decrease of 0.00038 in the parameter $p$. Despite the wide range of population density in our sample from 0.96 to 291.88, the impact of population density on the parameter $p$ is still trivial and negligible. However, we suggest several possible explanations for this association. First, more highly populated areas would receive more intense control interventions, with greater potential for behavioral changes. Second, the coefficient of population density was attenuated when variables negatively associated with the parameter $p$ were included in the model. The association weakens after controlling for other covariates, a similar effect to that obtained for the sociodemographic index. As such, it is possible that neither socio-demographic index nor population density captures what is important in this setting. Part of the spatial variation in the growth scaling may be due other factors not included in the model.

Our study is not exempt of limitations. First, in the absence of HIV infection data, we relied on AIDS incidence surveillance data as in prior studies (May and Anderson 1987; Colgate et al., 1989). While the natural history of HIV including the heterogeneity in incubation periods contribute to shaping the epidemic growth profiles, advancements in diagnostics and revisions to the AIDS case definition could artificially shorten the incubation period. For instance, during our study period (1982–1999), the AIDS case definition was revised once in 1997 (Oliver et al., 2014), but this only affect two states and at the very end of our study period. Second, previous studies have estimated AIDS underreporting at 42.7% in Brazil before 1999 (Bessa and Portela 1999). Third, socio-demographic variables were based on 1991 Census data. Finally, while our analyses accounted for variables relating to social-contact structures (i.e., population density, urbanicity, Gini index and SDI), we did not include more specific factors associated with HIV transmission (e.g. high-risk populations, multiplex relationships) (Rothenberg et al., 1998).

In summary, our findings indicate that the generalized growth model is convenient and advantageous in characterizing differences in epidemic growth patterns across different spatial scales. In the context of the HIV/AIDS epidemic in Brazil, substantial variation in epidemic growth scaling at the state level was associated with socio-demographic factors with Gini index and SDI negatively associated with the growth scaling parameter, whereas urbanicity was positively associated with this parameter. Our results should motivate further work to characterize differences in growth scaling across different disease systems and the factors driving their early transmission dynamics. From a public health point of view, our study could have significant implications by informing control interventions and public health resources allocation, e.g., target high urbanicity populations with epidemics characterized by higher epidemic growth scaling (e.g., near-exponential epidemic growth).

**Competing interest**

The authors have declared that no competing interests exist.

**Acknowledgements**

LD is supported through a 2CI Doctoral Fellowship at Georgia State University. GC acknowledge financial support from the NSF grant 1414374 as part of the joint NSF-NIH-USDA Ecology and Evolution of Infectious Diseases program; UK Biotechnology and Bio-
logical Sciences Research Council grant BB/M008894/1 and the Division of International Epidemiology and Population Studies at the Fogarty International Center, NIH.

Supplementary materials

References
Department of Informatics of the Brazilian Health System DATASUS.
The Brazilian Ministry of Health Systema de Informação de Agravos de Notificação (SINAN), http://portal.sinan.saude.gov.br.