Geographic area-based rate as a novel indicator to enhance research and precision intervention for more effective HIV/AIDS control

Xinguang Chen *, Kai Wang

Department of Epidemiology, University of Florida, Gainesville, FL, United States

ARTICLE INFO

ABSTRACT

Ending the HIV epidemic needs additional methods to better assess the incidence and prevalence of HIV infection. In this study, a new indicator — G-rate was developed for the evaluation of HIV epidemics across regions with regard to geographic area size. Different from the commonly used incidence and prevalence rates that assess the HIV epidemic with reference to population (termed as P rate in this study), G rate measures the number of new infections (incidence) or cases (prevalence) over a unit land area in one year. We demonstrated the utility of G rates using officially reported data on new HIV infections and persons living with HIV in the United States during 2000–2012. Findings of our analysis indicate that relative to P rates, G rates indicated a quicker increase in the HIV epidemic in the United States during the study period. In 2012, 4.6 persons were newly infected and 101.4 persons lived with HIV per 1000 km^2 land area. The five states with both highest P prevalence rates and highest G prevalence rates were Florida, Maryland, New York, New Jersey and Washington DC, which included New Jersey ranked 8th by P rate and excluded Massachusetts ranked 5th by G rate. In conclusion, adding G rates extends the conventional measurement system that consists of case count and P rate. Combining G rates with P rates provides a new approach for information extraction to support precision intervention strategy toward the goal of creating an AIDS-Free Generation.

1. Introduction

We have been living under the threat of the acquired immunodeficiency syndrome (AIDS) for more than three decades (UNAIDS) since the first reported AIDS case among homosexual men in 1981 in the United States (Gottlieb et al., 1981). AIDS is an infectious disease caused by the human immunodeficiency virus (HIV) (Alistar and Brandeau, 2012), and transmitted through sex, blood transfusion, and mother-child during pregnancy, child birth and breastfeeding. To end the HIV epidemic, the U.S. President's Emergency for AIDS Relief (PEPFAR), an initiative of the U.S. Government, has established a blueprint to create an AIDS-Free Generation (U.S. Department of State, 2012). In 2015, the United Nation adopted the strategy proposed by Jointed United Nations Programme on HIV/AIDS to end the AIDS epidemic by 2030 (UNAIDS, 2015). To achieve these great and ambitious goals, more comprehensive methods are needed to gauge the HIV/AIDS epidemic. Such methods should be able to add new information to better inform evidence-based decision making and to optimize resource allocation for best outcomes. Traditional epidemiologic measures provide data on total number of infections, incidence, prevalence, as well as time trends of these measures; however, no data are available about the epidemic with regard to geographic area size.

To control the HIV epidemic, it is important to know *the total number* of persons who have already been infected with HIV and who are newly infected. It provides basic data for public health decision-makers to estimate the amount of money and resources needed to successfully fight the epidemic. For example, data from World Health Organization (WHO) indicate that globally approximately 37 million persons live with HIV (PLWH), of whom 1.2 million are in the United States. If $1000 per PLWH per year is needed for HIV/AIDS control, a total of $37 billion must be budgeted worldwide and $1.2 billion for the United States alone. However, as an indicator of HIV epidemic, the total count of PLWH or new infections is inadequate. Although a larger number indicates a higher risk of HIV infection, the total count is affected by population size that varies dramatically across countries and regions. Given the same level of an epidemic, a country with a larger population will have a greater count of PLWH and new infections such as the United States than a country with a smaller population, such as Australia, Spain and Netherlands.

Prevalence and incidence are the two most commonly used measures in epidemiology. Prevalence measures the number of existing infections among a unit of at-risk population (e.g., 1000 or 100,000) in one year and incidence measures the number of new infections in the same period. Relative to total count, a population-based rate (termed as P rate

* Corresponding author at: Department of Epidemiology, University of Florida, 2004 Mowry Road, Gainesville, FL 32610, United States.
E-mail address: jimax.chen@ufl.edu (X. Chen).

thereafter), including *P prevalence* and *P incidence* provides new information regarding the HIV/AIDS epidemic. For example, data from the Centers for Disease Control and Prevention (CDC) indicated that 1.218 million PLWH in the United States in 2014, a very large number; but the P prevalence rate was only 18.5/100,000 (CDC, 2016); while data from UNAIDS indicated an estimate of 0.039 million in Botswana in 2014, a much smaller number; but the prevalence rate was as high as 25,200/100,000. In addition to describing levels of the HIV/AIDS epidemic, information from P rates can be used by decision-makers to prioritizing countries/places for intervention in order to achieve a pre-determined goal.

Inclusion of P rates strengthens the total count of persons living with HIV to inform decision-making and strategic planning for HIV control by factoring in population size; however, the size of geographic area where the population resides is ignored. Geographic area size plays a crucial role in the HIV epidemic. Given the same P rate, the likelihood is much greater for HIV to spread from one to another in a population residing in a crowed urban area than the same number of population residing in a large rural area (Sattenspiel, 2009). Simply mapping a P rate by geographic areas does not provide complete information about geographic differences of the HIV/AIDS epidemic, underscoring the need for new measures.

In this study, we reported our work attempting a new indicator by factoring in geographic area sizes. We demonstrated the new indicator using reported data on number of new infections as well as PLWH by states in the United States. Our purpose is to expand the current total count and P rate system in epidemiology by adding a geographic area-based measure.

2. Methods

2.1. Geographic area-based rate

We defined the geographic area-based rate (G rate) as the number of persons N with an event (e.g. newly infected or living with HIV) in one year within a jurisdiction (e.g. a district, a state, or a nation) over the total geographic area A of the jurisdiction:

$$G\ rate = N/A \tag{1}$$

This defined G rate was used to assess incidence and prevalence of HIV in the United States, overall and by the 50 states and the District of Columbia (DC). For each state, G incidence rate was computed as the number of newly infected HIV cases in a state during one year over the total area size A of the state:

G incidence rate
$$= Number\ of\ new\ cases\ detected\ in\ one\ year/A\left(1000\ km^2\right) \tag{2}$$

As Eq. (2) indicates, G incidence rate measures the number of new infections in a unit time over a unit geographic area, a higher G rate indicates more new infections in a jurisdiction. For convenience of reporting, two geographic units were used, 100 km² and 1000 km².

Likewise, the G prevalence rate of HIV was estimated as the number of PLWH in a state in one year over the total area size A of the state:

G prevalence rate
$$= Number\ of\ infected\ persons\ in\ one\ year/A\left(1000\ km^2\right) \tag{3}$$

As Eq. (3) indicates, a higher G prevalence rate means more PLWH in a unit time and area of a jurisdiction.

It is worth noting that the G rate we defined in this study is conceptually a measurement of geographic density of a disease. However, we elected not using the term *density rate* to avoid term duplication. The term incidence density was introduced in the 1980s by other researchers to measure the number of new disease cases for a population over a unit of time (but not area size) (Beaumont et al., 1985; Mutgi et al., 1988). This measure has been frequently used in sampling and epidemiological research (Greenland, 2013; Liu et al., 2015) although it has little to do with geographic density.

2.2. Population-based rates and other measures

To illustrate the significance of G rates, additional indicators included were: count of persons living with HIV and newly infected, overall and by states, and two population-based P rates that are commonly used in epidemiology:

P incidence rate
$$= Newly\ infected\ cases\ in\ a\ year/Population\ at\ risk(100,000) \tag{4}$$

P prevalence rate
$$= All\ infected\ persons\ in\ a\ year/Population\ at\ risk(100,000) \tag{5}$$

In addition to total count, P rate and G rate, we also examined another indicator by dividing G rate with population size and termed it as D rate. D rates measure the number of infected persons per unit population per geographic area. It can be considered as indicator not affected by population density. To focus on G rate in this study, we elected not to show the details about D Rate. Interested readers can investigate D rate by following our discussion on G rate in this study.

2.3. Sources of data

Data regarding the number of new HIV infections and number of PLWH (aged 13 and above) in the United States were derived from CDC's reports (CDC, 2008, 2016; U.S. Department of State, 2012; UNAIDS, 2015), overall from 2000 to 2012 and by single states for 2012. Data for annual population from 2000 to 2012 (aged 13 and above) and land area size (km²), overall and by states were derived from the US Census Bureau (2012). These data were directly downloaded from the websites and organized using the excel file for further analysis.

2.4. Statistical analysis

Data analysis was conducted in four steps. We first computed the G incidence rate and G prevalence rate using Eqs. (2) and (3) and P incidence and P prevalence using Eqs. (4) and (5). We then plotted the estimated incidence and prevalence of G rates and P rates respectively by year across the 2000–12 to compare the temporal trends. Followed the plotting step, we ranked and compared the 2012 G prevalence rates with P rates of individual states to illustrate the role of the new indicators in extracting additional information regarding the disease epidemiology. Lastly, we cross-plotted the ranks of G prevalence rates with those of P rates to illustrate the significance in combining the two indicators to better and more precisely informing public health planning and decision-making. Statistical analyses were conducted using the commercial software of MS Excel (Version 2010, Microsoft, Seattle, WA) and SAS version 9.4 (SAS Institute, Inc. Carry, NC).

3. Results

3.1. Time trends in the HIV epidemic

Fig. 1 depicts the time trends in the HIV epidemic in the United States during 2000–12 using both G rate and P rate. Results from panel A of Fig. 1 indicate that according to G rates, in 2000 there were 30.9 PLWH per 100 km² in the United States; and this number increased to 101.4 in 2012. Although both P rates and G rates captured the ups and downs of the HIV epidemic over time, G rates showed a quicker increase
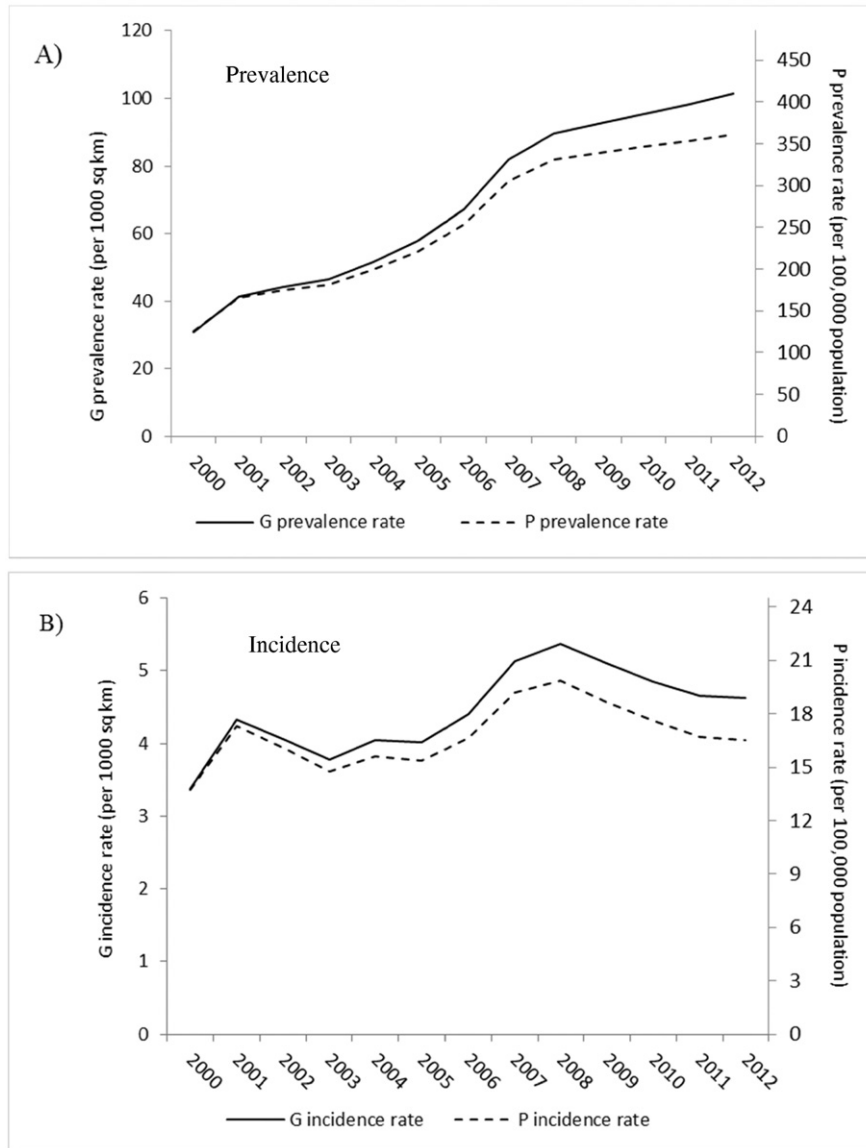
Fig. 1. G rates and P rates of HIV/AIDS epidemic in the United States (panel A: prevalence; panel B: incidence).

in the prevalence of HIV epidemic over time than P rates. Obviously, G rates provide new information regarding the HIV epidemic in the United States — P rate tells us the number of PLWH among 100,000 US population while G rates tells us the number of PLWH over the US land with an area size of 1000 km². Regardless of new infection, G rate is anticipated to increase as more PLWH live longer due to effective treatment.

Likewise, results in panel B of Fig. 1 indicate that new HIV infections in the United States increased quicker if G incidence rates were considered, relative to P incidence rates. Likewise, both indicators successfully captured the zigzags in time trends of new HIV infection; G incidence rates showed a quicker increase than the corresponding P rates. P rate tells the number of new infections among 100,000 US population while G rate tells the number of new infections over the US land with an area of 1000 km². Increases in G incidence rate indicates more new infections since there has been no change in the US land area during this period.

3.2. State differences in the HIV epidemic in 2012

Table 1 summarizes the information of the HIV epidemic for the 50 US states and DC using different measures. When G rates were

considered, Alaska ranked the lowest with 0.02 new infections per 1000 km² and 0.5 PLWH per 100 km²; while DC ranked the highest with 4189.9 new infections per 1000 km² and 13,734.2 PLWH per 100 km².

However, while P rates were considered, North Dakota ranked the lowest with 0.43 new infections and 56 PLWH per 100,000 persons; and DC remained to rank the highest with 120 new infections and 3936 PLWH per 100,000 persons.

3.3. Comparison of the ranks of G rates and P rates by states

All US states were ranked by the estimated G and P prevalence rates respectively (Fig. 2). Results in the figure indicate that the five states with the highest G rates were Massachusetts, New York, Maryland, New Jersey and DC. Among the five states, three remained as the highest when P rates were considered, they were DC, New York and Maryland; other two ranked lower by P rates with New Jersey moving to the 8th and Massachusetts to the 12th. The five states with the lowest G rates were Alaska, Wyoming, Montana, South Dakota and Idaho. These states remained as the lowest five by P rates, although their actual ranks changed.

Table 1
Land area, population, prevalence and incidence of HIV/AIDS measured using G rates and P rates respectively, 13 years of age and above, United States, 2012.

State	Total area[a]	Population	G rate		P Rate	
			Incidence[b] (Rank)	Prevalence[c] (Rank)	Incidence[d] (Rank)	Prevalence[e] (Rank)
Alabama	131,171	4,022,346	4.88 (23)	10.98 (24)	15.91 (18)	358 (20)
Alaska	1,477,953	593,985	0.02 (51)	0.05 (51)	4.71 (36)	133 (44)
Arizona	294,207	5,382,060	2.16 (29)	5.51 (30)	11.80 (24)	301 (22)
Arkansas	134,771	2,436,975	1.30 (35)	4.30 (33)	7.18 (32)	238 (29)
California	403,466	31,440,823	12.10 (12)	45.43 (10)	15.53 (19)	583 (7)
Colorado	268,431	4,285,714	1.44 (34)	4.69 (31)	9.01 (30)	294 (24)
Connecticut	12,542	3,040,541	23.76 (9)	107.64 (6)	9.80 (27)	444 (13)
Delaware	5047	769,231	28.33 (8)	85.20 (9)	18.59 (15)	559 (9)
D.C.	158	551,321	4189.87 (1)	13,734.18 (1)	120.08 (3)	3936 (1)
Florida	138,887	16,460,746	35.42 (4)	92.09 (8)	29.89 (11)	777 (4)
Georgia	148,959	8,116,147	13.96 (10)	38.47 (11)	25.63 (12)	706 (5)
Hawaii	16,635	1,166,667	4.27 (24)	21.04 (18)	6.09 (35)	300 (23)
Idaho	214,045	1,279,070	0.15 (46)	0.51 (50)	2.50 (42)	86 (47)
Illinois	143,793	10,702,576	11.99 (13)	31.78 (13)	16.11 (17)	427 (14)
Indiana	92,789	5,402,844	5.41 (22)	12.29 (22)	9.29 (28)	211 (34)
Iowa	144,669	2,545,455	0.82 (39)	1.94 (41)	4.64 (37)	110 (46)
Kansas	211,754	2,356,688	0.72 (40)	1.75 (42)	6.49 (34)	157 (37)
Kentucky	102,269	3,640,351	3.72 (26)	8.12 (27)	10.44 (26)	228 (31)
Louisiana	111,898	3,791,946	10.43 (15)	20.20 (19)	30.78 (10)	596 (6)
Maine	79,883	4,022,346	0.60 (41)	2.25 (40)	1.19 (44)	157 (38)
Maryland	25,142	593,985	53.34 (3)	172.22 (3)	225.76 (1)	880 (3)
Massachusetts	20,202	5,382,060	33.51 (5)	133.65 (5)	12.58 (23)	477 (12)
Michigan	146,435	2,436,975	5.59 (21)	11.95 (23)	33.61 (9)	211 (33)
Minnesota	206,232	31,440,823	1.55 (33)	4.07 (34)	1.01 (46)	188 (36)
Mississippi	121,531	4,285,714	3.88 (25)	8.48 (26)	11.01 (25)	420 (16)
Missouri	178,040	3,040,541	3.02 (27)	7.41 (28)	17.66 (16)	263 (27)
Montana	376,962	769,231	0.06 (48)	0.17 (49)	2.86 (41)	77 (48)
Nebraska	198,974	551,321	0.39 (43)	1.11 (45)	13.97 (20)	145 (40)
Nevada	284,332	16,460,746	1.27 (37)	3.38 (38)	2.20 (43)	421 (15)
New Hampshire	23,187	8,116,147	2.16 (29)	6.90 (29)	0.62 (49)	141 (41)
New Jersey	19,047	1,166,667	70.62 (2)	226.28 (2)	115.29 (4)	580 (8)
New Mexico	314,161	1,279,070	0.38 (44)	1.15 (44)	9.23 (29)	210 (35)
New York	122,057	10,702,576	31.40 (6)	145.01 (4)	35.80 (7)	1070 (2)
N. Carolina	125,920	5,402,844	10.95 (14)	25.41 (14)	25.52 (13)	395 (17)
N. Dakota	178,711	2,545,455	0.06 (48)	0.18 (48)	0.43 (51)	56 (51)
Ohio	105,829	2,356,688	10.30 (16)	21.64 (17)	46.25 (6)	237 (30)
Oklahoma	177,660	3,640,351	1.63 (31)	3.77 (35)	7.97 (31)	214 (32)
Oregon	248,608	3,791,946	1.03 (38)	3.38 (39)	6.72 (33)	256 (28)
Pennsylvania	115,883	4,022,346	12.31 (11)	35.29 (12)	35.45 (8)	378 (18)
Rhode Island	2678	593,985	28.75 (7)	93.35 (7)	12.96 (22)	278 (25)
S. Carolina	77,857	5,382,060	9.30 (17)	24.79 (15)	13.45 (21)	489 (11)
S. Dakota	196,350	2,436,975	0.14 (47)	0.26 (47)	1.11 (45)	76 (49)
Tennessee	106,798	31,440,823	8.41 (19)	17.98 (20)	2.86 (40)	357 (21)
Texas	676,587	4,285,714	6.24 (20)	15.42 (21)	98.47 (5)	497 (10)
Utah	212,818	3,040,541	0.52 (42)	1.36 (43)	3.65 (38)	132 (45)
Vermont	23,871	769,231	0.25 (45)	3.39 (37)	0.78 (48)	150 (39)
Virginia	102,279	551,321	9.22 (18)	24.54 (16)	171.04 (2)	367 (19)
Washington	172,119	16,460,746	2.92 (28)	8.95 (25)	3.05 (39)	268 (26)
W. Virginia	62,259	8,116,147	1.30 (36)	3.53 (36)	1.00 (47)	139 (42)
Wisconsin	140,268	1,166,667	1.63 (31)	4.56 (32)	19.63 (14)	134 (43)
Wyoming	251,470	1,279,070	0.03 (50)	0.13 (50)	0.55 (50)	67 (50)
Total	9,147,594	295,516,598	4.51	13.41	13.97	415

Note: HIV/AIDS: human immunodeficiency virus/acquired immune deficiency syndrome; G rate: geographic area-based rate; P rate: population-based rate.

[a] In square kilometers.

[b] G incidence rate = number of newly infected persons in 2012 per 1000 km^2.

[c] G prevalence rate = number of persons living with HIV in 2012 per 100 km^2.

[d] P incidence rate = number of newly infected persons in 2012 per 100,000 population.

[e] P prevalence rate = number of persons living with HIV in 2012 per 100,000 population.

The differences in ranking between G rates and P rates of HIV across individual states suggest the significance of G rates to provide a new piece of information to characterize the HIV epidemic across different states within a country. Since the ranking for both G and P incidence was similar to that of the prevalence, the results were not reported here and are available from the authors upon request.

3.4. Integrating G rates with P rates

Fig. 3 presents the cross-plot of the ranks of G rates with those of P rates for all 50 US states plus DC. The ranks were reversed such that a lower value indicating a higher prevalence rate. The plot area was further divided into four quadrants using rank 25 as the cutoff point. Clockwise, the first quadrant contains four states (i.e., Nevada, Colorado, Arizona and Mississippi). These states were characterized by high P rates and low G rates, which epidemiologically implies that although many people were infected with HIV *only* if population size was considered; not many were infected when the land area was considered. Consequently, despite high P rates, rapid HIV spread was less likely because the infected persons were sparsely distributed across a large geographic area.

The second quadrant contains 23 states, and typically ones include Wyoming, Oklahoma, and Missouri. These states were characterized by both low G rates and low P rates. With regard to HIV spread, these
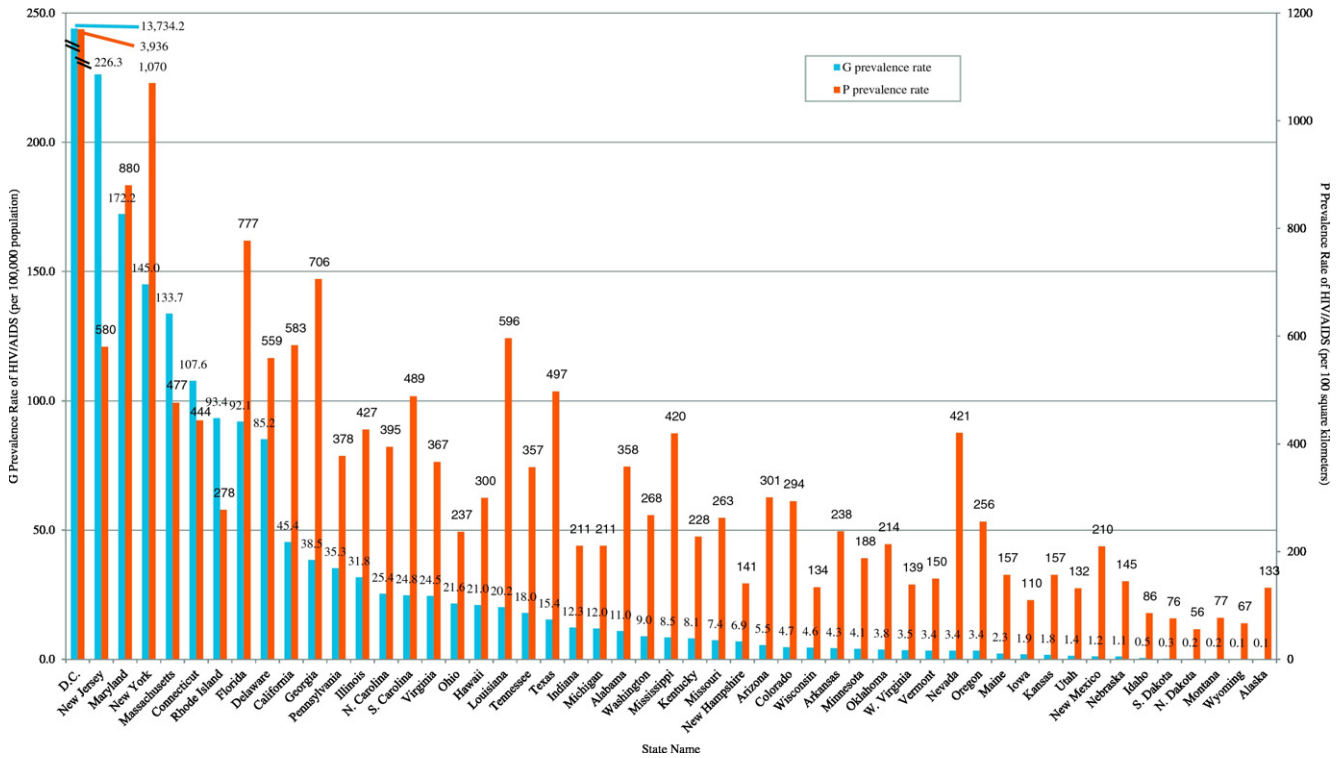
Fig. 2. HIV/AIDS G and P prevalence rates of the 50 US states and DC, all populations ≥13 years old, 2012.

23 states are at the lowest risk. However, it will cost much to reach individual infected persons for HIV/AIDS control because of the low case density across geographic area and population.

The third quadrant contains five states and they are Michigan, Indiana, Ohio, Washington and Rhode Island. The HIV epidemic in these states was characterized by high G rates and low P rates, suggesting higher risk than states in the second quadrant to have rapid HIV spread. Although not many persons were infected considering the population size, the infected persons are densely distributed over a geographic area, increasing the chance for the virus to spread from one to another through easy inter-personal contact.

All states in the fourth quadrant were at the highest risk for HIV spread because of both high G rates and P rates. This quadrant depicts a scenario with many HIV infected persons living in highly crowed areas, and the most typical states were Washington DC, New York, New Jersey and Maryland.

The plotted area was also divided, using ranks 20, 30 and 60 as radians into four regions respectively. The most inner band region contained 13 states with both high G rates and high P rates, with DC, Maryland, New York and New Jersey at the inner core. These states should be considered as the top priority for HIV treatment and control. In contrast, the farthest band region consisted of nine states, such as Alaska, Montana, Wyoming, North Dakota and South Dakota. Relative to states in other bands, the burden for HIV treatment and control would be the smallest. Likewise, the risk of HIV spread and the burden for treatment and control for the states located in the two middle brands were moderate comparing to those located in the most inner and the farthest band regions.

4. Discussion and conclusion

In this study, we reported our work in developing a new indicator to assess the HIV epidemic by taking into account the size of land area. This research was based on decades of our epidemiological research and teaching and intention to promote the effort toward an AIDS-Free Generation as proposed by the US government and the goal to end the AIDS epidemic by 2030 set by USAIDS. Conventionally, in addition to total count, only population-based P rates are used in assessing the epidemic of a disease, including HIV/AIDS. While P rates make it possible to compare disease epidemic across countries and regions with different population size; it cannot be used to make such comparison when land area differs by regions within a country and by countries across the globe. In this case, the G rate we proposed provides a useful option. We demonstrated this new measurement method and showed some of its applications in assessing the HIV epidemic with real data.

4.1. G rate as an innovative indicator

To the best of our knowledge, we are the first to introduce G rate to extend the current disease measurement systems in modern epidemiology. The size of geographic area has been considered as an important measurement component in other fields, including environmental science (Levin and Kerster, 1969; Nambinina et al., 2016), demography (Langford and Unwin, 1994; Pearl and Parker, 1922), and economics (Bickenbach et al., 2016; Gallup et al., 1999). However, no reported studies have ever factored this component into any formal rate indicators in epidemiology. Adding G rate makes the epidemiological measurement systems more comprehensive. For any disease in a population, the total count of patients provides a direct measure of disease burden, P rate assesses the number of patients per unit population, while G rate adds information regarding the number of patients per unit geographic area, forming an integrative measurement system. No epidemiology of a disease would be complete without inclusion of information derived from G rate.

4.2. Significant characteristics of G rate

Findings of our study indicate that G rates possess significant characteristics for epidemiologic research and practice. First of all, G rates for a disease tell us how many people have already been infected in a given geographic area in a year. Regardless of disease type and complexity of the epidemics, a higher G rate means more people are sick in a
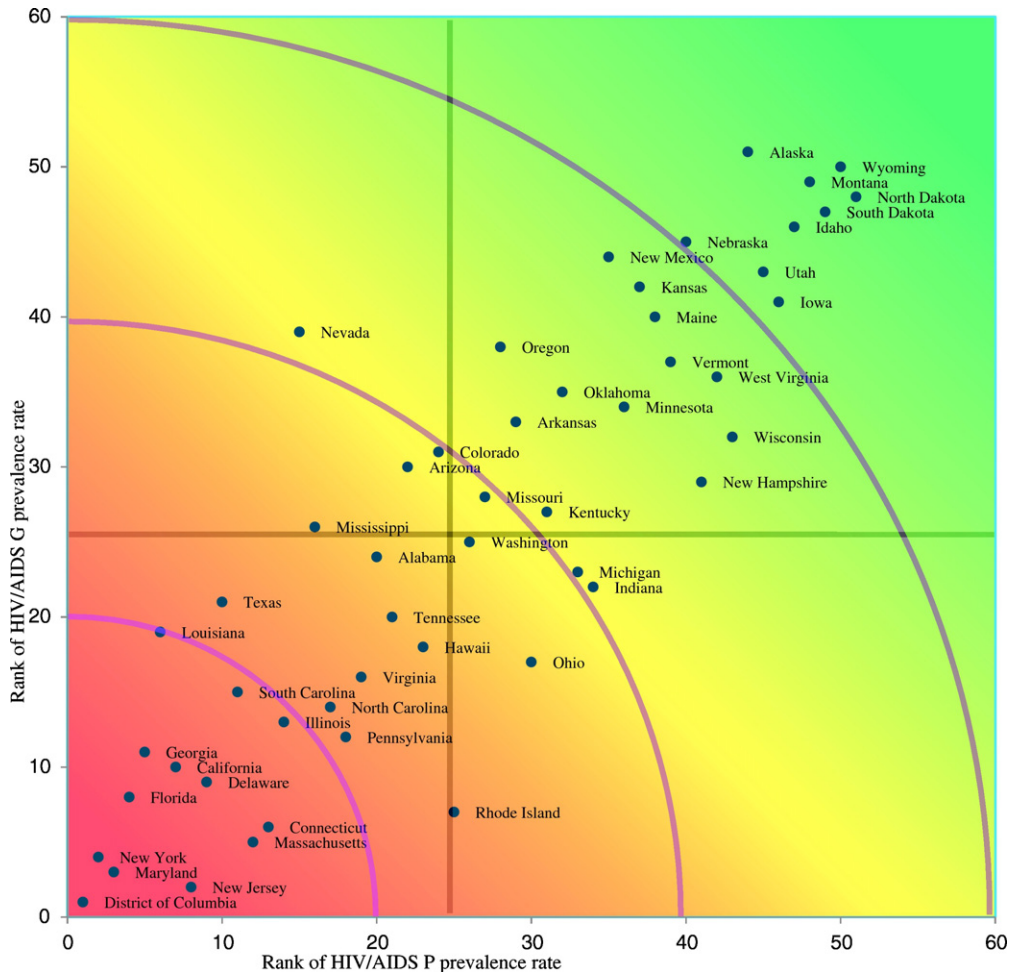
Fig. 3. Cross-plot of the reverse ranks of G rates with P rates of HIV prevalence of US States, 2012.

given geographic area. For infectious disease, this means more contact opportunities between the infected and the non-infected; while for non-infectious diseases, it provides information regarding geographic clustering complementary to P rate. Second, G rates of a disease are not affected by the area size, therefore these rates can be used to compare disease epidemic across jurisdictions within a country and across countries over the globe. Third and more importantly, as we demonstrated in this study, a combination of the G rate we proposed in this study with the classic P rate is superior to either one alone to inform planning and decision making for disease prevention and control. Places with high G rates and high P rates are at the highest risk for infectious disease transmission, and must be considered as the top priority for control, including prevention and treatment. In this study, we demonstrated the significance using HIV/AIDS as an infectious disease example, further research is suggested to investigate the role of G rate in studying non-infectious diseases.

4.3. Recommendation for HIV/AIDS control in the United States

According to the 2012 data, five states with the highest P prevalence rate were Georgia, Florida, Maryland, New York and Washington DC; and five states with highest G prevalence rates were Massachusetts, New York, Maryland, New Jersey, and Washington D.C. However, when P rates and G rates both were considered, Washington DC, Maryland, New York, New Jersey, and Florida were top five. An optimal strategy for HIV control in the US would be to allocate proportionately more resources to these five states. Among the five states, Washington DC, Maryland, New York remained within the top five regardless the use

of P rate or G rate. However, the final five states included New Jersey not within the top five by P rate but excluded Georgia within the top five by P rate. New Jersey with a P prevalence rate of 580/100,000 was ranked top 8 but its G rate was ranked the 2nd, indicating 2.26 HIV positive persons per square kilometer in the state. The high density of HIV positive persons suggests both high risk for HIV spread and more resources needed for intervention (International, 2012; Mann and Tarantola, 1996; Morens et al., 2004; Quinn, 1996).

There are limitations to the present study. First, we tested the G rate method using only the US data at the national level from 2000 to 2012 and state level in 2012. More research is needed to test the method using additional data from other time for other diseases. Second, just like P rate that does not consider differences in population age structure, G rate does not consider urban, suburban, and rural differences in residential arrangement and other variations in population density. Caution should be used when results measured with G rates are interpreted in epidemiological research and public health practice. D rate can be considered as a standardized G rate because the denominator of D rate in fact has incorporated factors of population density.

Despite the limitations, the G rate we proposed in this study adds a new tool, in a timely manner to advance HIV epidemiology and to provide new information supporting HIV/AIDS control in the United States and across the globe.

Funding

Conflicts of interest

None.

References

Alistar, S.S., Brandeau, M.L., 2012. Decision making for HIV prevention and treatment scale up: bridging the gap between theory and practice. Med. Decis. Mak. 32, 105–117.

Beaumont, J., Steenland, K., Minton, A., 1985. Selection of controls in nested case-control studies — a computer-program for incidence density sampling. Am. J. Epidemiol. 122, 510.

Bickenbach, F., Bode, E., Nunnenkamp, P., Söder, M., 2016. Night lights and regional GDP. Rev. World Econ. 152, 425–447.

U.S. Census Bureau, 2012. 2010 Census of Population and Housing.

CDC, 2008. HIV/AIDS Surveillance Report, Cases of HIV Infection and AIDS in the United States and Dependent Areas, 2006. p. 18.

CDC, 2016. HIV/AIDS in the United States by Geographic Distribution.

Gallup, J.L., Sachs, J.D., Mellinger, A.D., 1999. Geography and economic development. Int. Reg. Sci. Rev. 22, 179–232.

Gottlieb, M.S., Schroff, R., Schanker, H.M., et al., 1981. *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N. Engl. J. Med. 305, 1425–1431.

Greenland, S., 2013. Confounding of incidence density ratio in case-control studies. Epidemiology 24, 624–625.

International, PA, 2012. Why Population Matters to Infectious Diseases and HIV/AID.

Langford, M., Unwin, D.J., 1994. Generating and mapping population density surfaces within a geographical information system. Cartogr. J. 31, 21–26.

Levin, D.A., Kerster, H.W., 1969. The dependence of bee-mediated pollen and gene dispersal upon plant density. Evolution 560–571.

Liu, L., Wu, J.Y., Yue, S., et al., 2015. Incidence density and risk factors of diabetic retinopathy within type 2 diabetes: a five-year cohort study in China (report 1). Int. J. Environ. Res. Public Health 12, 7899–7909.

Mann, J.M., Tarantola, D., 1996. AIDS in the World Two. Oxford University Press on Demand.

Morens, D.M., Folkers, G.K., Fauci, A.S., 2004. The challenge of emerging and re-emerging infectious diseases. Nature 430, 242–249.

Mutgi, A.B., Zeitler, R.R., Helms, C.M., 1988. Pneumonia in renal-transplant recipients — incidence density and temporal trend. Clin. Res. 36, A348-A48.

Nambinina, F.A., Champolivier, L., Debaeke, P., Maury, P., 2016. Source and sink indicators for determining nitrogen, plant density and genotype effects on oil and protein contents in sunflower achenes. Field Crop Res. 192, 33–41.

Pearl, R., Parker, S.L., 1922. On the influence of density of population upon the rate of reproduction in *Drosophila*. Proc. Natl. Acad. Sci. 8, 212–219.

Quinn, T.C., 1996. Global burden of the HIV pandemic. Lancet 348, 99–106.

Sattenspiel, L., 2009. The Geographic Spread of Infectious Diseases: Models and Applications. Princeton University Press.

U.S. Department of State, 2012. PEPFAR Blueprint: Creating an AIDS-Free Generation. The Office of the Global AIDS Coordinator.

UNAIDS, 2015. One the Fast Track to End AIDS: 2016–2021 Strategy. UNAIDS, Geneva.