Original Article

Population Dependency of Measles, Syphilis, and Amebiasis in Japan and Community Evolution

Hiroshi Yoshikura^{1*} and Fumihiko Takeuchi²

¹National Institute of Infectious Diseases, Tokyo 162-8640; and

²Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 162-8655, Japan

SUMMARY: The population dependency of measles, syphilis, and amebiasis was expressed as $P = kN^m$, where P , N , and m were number of patients, population size, and a constant (~ 2 for measles, and $1.3\sim 1.4$ for syphilis and amebiasis), respectively. The population size dependency emerged only when conditions other than population size, such as infant mortality, hygienic condition, vaccination practices, and others, improved to the same level in all the prefectures in Japan. The formation of prefectures and municipalities was well simulated by the random coin toss assuming that people are attracted to a community with a probability proportional to the number of the residents to the 1.3rd power. When the number of inflow population was plotted against the number of the resident population in a prefecture, or when the number of coins that were added in a round was plotted against the number of coins that were present before the coin toss, the plots fell on a straight line with the slope ~ 1.3 , which was almost the same as the slope obtained when the number of cases of syphilis or amebiasis was plotted against the population size.

INTRODUCTION

Since 2008, cases of measles, rubella, human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), and syphilis were population size-dependent in Japan; the number of patients (P) was related to the population size (N) with the equation $P = kN^m$, where m was ~ 2 for measles and rubella (1) and ~ 1.5 for HIV/AIDS (2). In this study, we examined the data before 2008 to assess the possible influence of social conditions on this phenomenon. In addition, we tried to answer the question why m was $1.3\sim 1.5$ for HIV/AIDS and other sexually transmitted diseases (3).

MATERIALS AND METHODS

The following sources of data were used: i) for the infectious disease statistics, Infectious Diseases Statistics under the Infectious Disease Prevention Act (<http://www.mhlw.go.jp/toukei/list/densenbyou.html>) and National Epidemiological Surveillance of Infectious Diseases (NESID) (<http://www.nih.go.jp/niid/ja/all-surveillance/2270-idwr/nenpou/3363-foward2011.html>), and ii) for the population data, the Census (<http://www.e-stat.go.jp/SG1/estat/GL02100104.do?tocd=00200521>,

http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&listID=000001135862&requestSender=dsearchhttp://uub.jp/rnk/cktv_j.html, and https://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&listID=000001101885&requestSender=estat). All the analyses were performed using Microsoft Excel 2010.

Patient-population plot: The plot of the number of patients per prefecture in the y-axis versus the population size of prefectures in the x-axis (1). The approximation was the power approximation (for the convenience of the graphics, “0,” which was rare, was replaced by “0.1”).

Zipf type plot: The population size was plotted on the x-axis, and the ranking number starting with the largest was plotted on the y-axis (3). For simulated data, the number of coins in the boxes was plotted on the x-axis and the ranking number of the boxes was plotted on the y-axis.

Simulating evolution of the community population by the random coin toss: Let the total number of the boxes = m . The probability of receiving a coin in a box was set proportional to the number of coins to the 1.3 power (the best fit among $1.1\sim 1.4$ power). If at the i step box N_k had received N_{ki} coins, the probability of receiving the next coin by the box was $N_{ki}^{1.3} / (N_{i1}^{1.3} + N_{i2}^{1.3} + \dots + N_{ik}^{1.3} + \dots + N_{im}^{1.3})$. For the prefectures excluding Okinawa ($m = 46$), after a total of 1,000 tosses (the initial even distribution of 10 coins/box followed by the 540 tosses, according to the above rule), the boxes were ranked in descending order by the number of the coins in the boxes. The number of coins received in each box at each step was counted. For the municipalities (m

Received May 11, 2016. Accepted August 29, 2016.
J-STAGE Advance Publication October 31, 2016.

DOI: 10.7883/yoken.JJID.2016.222

*Corresponding Author: Mailing address: Department of Food Safety, Ministry of Health, Labour and Welfare, 1-2-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-8916, Japan.
E-mail: yoshikura-hiroshi@mhlw.go.jp

= 1,700), after total 20,000 tosses (the initial even distribution of 10 coins/box, followed by the 3,000 tosses according to the rule), the boxes were ranked in descending order by the number of coins in the boxes. The number of coins received in each box at each step was counted. The simulation was carried out by using the R software (<https://cran.r-project.org/>).

RESULTS

Measles: Under the Infectious Disease Prevention Act of 1954, which was effective until 1999, measles was reportable but not notifiable. In 1999, the infectious disease surveillance was placed under the NESID. Until 2007, cases of measles were reported from ~3,000 pediatric sentinel clinics and from approximately 450 core sentinel hospitals for adults (4, 5). Between 2007 and 2008, the reporting of measles cases in Japan was changed from sentinel reporting to the reporting of all measles cases.

The annual incidences under the Infectious Disease Prevention Act and those under the NESID system, respectively, are shown in Figs. 1-A1 and 1-A2. The patient-population plot is shown in Fig. 1B. If the number of patients is proportional to the population size, the slope is one. In 1948~1974, the slope was 0.45~0.32, indicating a higher incidence of measles in the less populated rural prefectures. The slope became steeper from 2000 to 2008. The change in the reporting system from sentinel reporting to the reporting of all cases between 2007 and 2008 did not influence the slope, as it was 1.71 and 1.88 in these respective years. In Fig. 1C, the graph of the slopes ($\circ$) and correlation coefficient (CC) ($\triangle$) between the number of patients and the population size are shown. The slope became ~ 2 in 2007 and the correlation CC became ≥ 0.7 in 2005, which was maintained thereafter (Fig. 1D)(1). With the shift from sentinel reporting to all cases reported, the percentage of the ≥ 15 year cases jumped from approximately 10% to 60%, indicating that cases of ≥ 15 years had already rep-

Fig. 1. Measles 1948–2013. A1 and A2: Number of measles cases (× 1,000) per year in 1948 – 1999 and in 2000–2013, respectively; the plots later than 2008 when reporting of all the cases started are shown by shaded symbols (A2, C, and D). B: Plot of patients/prefecture (y-axis) vs. population (× 1,000)/prefecture (x-axis). C: The slopes of the patient-population plots from 1948 to 2008 ($\circ$) and the correlation coefficient (CC) of number of patients and the population size of prefectures ($\triangle$). D: Percentage of patients ≥ 15 years among the total patients. E: Infant mortality (< 1 year) per 1,000 births versus population (× 1,000,000)/prefecture.

resented a significant proportion of cases before 2008 (6).

In 1970, when the infant mortality rate was above one in 100,000, the slope of the infant mortality-population plot was -0.18 (i.e., the infant mortality rate was higher in the less populated rural prefectures). In 2000, when the infant mortality rate decreased to 3 in 1,000,000, the slope became flat (-0.018) (Fig. 1E). In 1978, the measles vaccine was incorporated into the routine vaccination schedule (7), and in 2007, the measles elimination program was launched (8). Thus, the population dependency of the measles epidemics only emerged when public health and other conditions improved to nearly the same level in all the prefectures in Japan.

Syphilis and tuberculosis: The incidence of syphilis and tuberculosis decreased continuously from the mid-1940s to 2008 (Fig. 2A). The patient-population plots for tuberculosis and syphilis, respectively, are shown in Figs. 2B and 2C. The slopes obtained from Figs. 2B and 2C were plotted in Fig. 2D. The slope of syphilis ($\circ$) remained ~ 1.1 until 1970, when the slope suddenly shifted to ~ 1.4 . The shift corresponded to the rise of gross do-

mestic product of Japan triggered by the Tokyo Olympic Game in 1964 and flow of young people into the urban prefectures (Fig. 2E).

The slope of tuberculosis was ~ 1.4 in mid-1940s reflecting the poor health condition in the urban prefectures during the post-World War II period. It gradually decreased to $0.8\sim 1.0$ owing to the improved living conditions in the urban prefectures on one hand and to the increase of elderly population ($\diamond, \triangle$, and $\square$ in Fig. 2F) with higher rates of tuberculosis ($\bullet$ in Fig. 2F) (9) in the less populated rural prefectures on the other hand.

Bacillary dysentery: Bacillary dysentery is transmitted via food and water as well as from person to person. The incidence decreased from mid-1960s to mid-1970s, and then remained in the similar level (large triangles) (Fig. 3A). The incidence for age groups ≥ 15 years and < 15 years were comparable initially. From 1972 on, however, the incidence of the < 15 years continuously decreased while that of ≥ 15 years remained unchanged, which was probably due to the increase of the infection abroad of the adults (61% and 77% in 1993 and 1994)

Fig. 2. Syphilis and tuberculosis 1944–2012. A: Number of patients per population ($\times 1,000$)/year; $\circ$: syphilis; $\bullet$: tuberculosis. B and C: Patient-population plot of tuberculosis and syphilis, respectively. D: The slopes of the patient-population plots from 1948 to 2008 of syphilis ($\circ$) and that of tuberculosis ($\bullet$). E: % population 30–34 years in the y-axis and the ranking number of prefectures according to the population size. F: % population ≥ 65 years (1960: $\diamond$; 1980: $\triangle$; 2001: $\square$) and % tuberculosis patients ≥ 60 years among the all tuberculosis patients in 1987 ($\bullet$) in the y-axis and the ranking number in the x-axis.

(10). The slope of the patient-population plots was 1.3~1.4 from mid-1950s to mid-1960s, decreased to 0.6~1.0 until the mid-1970s, increased to 1.4~1.6 from the mid-1970s to the mid-1990s, and returned to ~1.0 in mid-2000s (Figs. 3C and 3E). The steep slope of ~1.4 that was observed until 1965 may reflect the insufficient sanitary conditions in the populated urban prefectures (11). The slope < 1.0 in 1970~1975 was attributable to the earlier improvement of living conditions in the urban prefectures. The slope > 1.0 in 1978~1998 (attaining 1.65 in 1993) was probably due to the frequent travel abroad of the urban residents. It became less steep later than 2000 probably because people went abroad irrespectively of their places of residence.

Amebiasis: Amebiasis spreads as a foodborne and/or waterborne infection as well as a sexually transmitted disease. The incidence decreased to its lowest levels in the mid-1970 (Fig. 3A), when the male cases (shaded

large circles in Fig. 3A) and the age group ≥15 years (closed small circles in Fig. 3B) started increasing, and the slope of patient-population plot became ~1.47 in 2000 (Fig. 3D and Fig. 3E); at the same time, the incidence became approximately 10-fold higher for men than for women (Fig. 3A), and patients aged < 15 years were negligibly few (Fig. 3B). This shift was probably linked to the change of the principal transmission route from foodborne illness to sexual transmission among young men (12).

Enterohemorrhagic *Escherichia coli* (EHEC): As shown in Fig. 3E, the slopes of the patient-population plot of EHEC, which is transmitted mainly through consumption of contaminated foods and occasionally from person to person, were plotted (13). The slopes of EHEC (shaded squares) were within the range of 0.7~1.0 (Fig. 3E).

From these considerations, it was deemed that orally

Fig. 3. Bacillary dysentery and amebiasis 1953–2014. A: Number of patients per year for bacillary dysentery (large shaded triangles: total; small open triangles: males; small closed triangle: females), amebiasis (large shaded circles: total; small open circles: males; small closed circles: females) and EHEC (squares). B: Number of bacillary dysentery /prefecture (triangles) and amebiasis (circles) classified by age (▲●: ≥ 15 years; △○: < 15 years). C and D: The patient-population plot of bacillary dysentery and amebiasis, respectively. E: Plot of slopes and CC between the number of the patients and the population sizes among prefecture for amebiasis (○), bacillary dysentery (▲), and EHEC (shaded squares).

transmitted amebiasis (before 2000), bacillary dysentery, EHEC, and tuberculosis were basically population size-independent and measles, amebiasis post-2000, and syphilis were intrinsically population-dependent, which was expressed by an equation $P = kN^m$, where m was ~ 2 for measles and $1.3\sim 1.4$ for amebiasis and syphilis. As $m = 2$ for measles was already explained (1), we asked why m was $1.3\sim 1.5$ for the sexually transmitted diseases including amebiasis, syphilis, and HIV/AIDS (2).

Evolution of community population and population dependency of infections: The Zipf type plot of the prefecture populations is shown in Fig. 4-A1. The prefectures were first arranged in the descending order of their population size in 1920 and ranked into categories of high (15 prefectures), middle (15 prefectures), and low (17 prefectures). In 1936, there were already 6 prefectures with an exceptionally large population size (i.e., Tokyo, Osaka, Hokkaido, Hyogo, Fukuoka, and

Aichi). The population size of these prefectures expanded dramatically between 1960 and 1970 (arrow, Fig. 4-A1). This period coincided with the expansion of the Japanese economy and emergence of the population-dependency of syphilis (Fig. 2D). The evolution of the individual prefectures is shown in Fig. 4-B1. The x-axis is the ranking number of prefectures in 1920, and the y-axis is the population size. The large prefectures on the left side, continuously expanded, while the smaller prefectures on the right side remained small. Some prefectures in the center of the graph, which were prefectures in the commuter belt of Tokyo, expanded from 1970 to 2008.

The evolution of the prefecture population was simulated using a hypothesis that the probability of a person to join a community was proportional to the number of residents to the 1.3 power. The Zipf type plot of the simulated population (Fig. 4-A2) closely resembled the real distribution (Fig. 4-A1), including the rightward

Fig. 4. (Color online) Evolution of community population size; the upper and the lower panels respectively show the real data and the simulations. A1: Zipf type plot of prefecture population. The prefectures with the high middle and low population sizes in 1920 are marked brown, green, and white. A2: Zipf type plot of the simulated community population sizes (1 of 3 independent trials, which resembled Fig. 4-A1 most). B1: Evolution of the individual prefectures. B2: Simulation of B1. C1.1: Zipf type plot of all the municipality population in Japan in 2010. C1.2: Zipf type plot of the municipality population for some prefectures in 2010. C2: Zipf type plot for the simulated 17,000 municipalities at the different steps of the coin toss.

protrusion of highly ranked boxes (i.e., ranking number 1, 2, 3, and so on) (arrows). The simulated data used for Fig. 4-A2 was similarly aligned with the actual data (Fig. 4-B1) to produce Fig. 4-B2. The simulation closely resembled the real plot, particularly in that some boxes that received initially smaller number of coins formed visible peaks after chance events during steps 20000~40000.

The Zipf-type plot of the municipality population in Japan and of the municipality population in the different prefectures are shown in Fig. 4-C1.1 and Fig. 4-C1.2, respectively. The curve for all of Japan and those for individual prefectures were similar in shape, indicating that the community size distribution was fractal (14). The simulation of the evolution of 1,700 individual municipalities, using the same algorithm that was used in simulating the prefectures, is shown in Fig. 4-C2. The simulated plots were quite similar to the plot for the municipality population in Japan (Fig. 4-C1.1), with a sharp rise on the left end and a sharp drop on the right end.

As shown in Fig. 5A, the number of inflow persons per prefecture in the y-axis is plotted against the population size of the prefectures in the x-axis. The plot was rightward ascending, and the slope was 1.23 for Japanese inflow population and 1.44 for non-Japanese inflow population. Similar plots for the simulated prefecture population are shown in Fig. 5B. The increment of the population between the 2 steps was plotted on the y-axis, and the population in the preceding step on the x-axis. All the plots fell on straight lines with slopes of 1.3~1.38. The slope of the plots of inflow population in

the real statistics (Fig. 5A) and the increment of the population in the simulation (Fig. 5B) were similar to that of syphilis, amebiasis, and HIV/AIDS (2), which was 1.3~1.5. When the number of inflow persons to municipalities was plotted against the number of the resident population, the slope was 1.06 for Japanese and 1.22 for non-Japanese (Fig. 5C). Using the data from Fig. 4-C2, the increment of the coins between the step 80000 and step 90000 was plotted on the y-axis and the number of coins in the boxes at the step 80000 on the x-axis (Fig. 5D). The slope was 1.27. The slope for non-Japanese inflow population (1.22) and the simulation (1.27) were similar and the slope of HIV/AIDS, syphilis, and amebiasis (1.3~1.4).

One observation that did not fit well to the other data was the slope for Japanese inflow population, which was much less steep than the others were. Though the reason is not clear, it could be due to the fact that the current municipalities were the consequence of the repeated re-grouping of communities to meet the governance needs (9,868 municipalities in Japan in 1953 were reduced to 1,718 in 2014) (15).

DISCUSSION

The hypothesis that the infection chance was proportional to the number of random encounter of persons well explained the relation between the number of measles patients (P) and the population size (N), which was $P = kN^2$ (1). However, for the multiplier 1.3~1.4 of the equation for the sexually transmitted diseases, $P =$

Fig. 5. Inflow population size distribution vs. resident population size distribution. A: The number of inflow population per prefecture (○: Japanese;●: non-Japanese) was plotted in the y-axis and the number of resident population per prefecture in the x-axis. B: Simulation of the panel A; the number of coins received during the last round of coin tosses was plotted in the y-axis and the number of coins before the start of the last round was plotted in the x-axis. C: The number of inflow population per municipality (○: Japanese;●: non-Japanese) was plotted in the y-axis and the number of resident population per municipality in the x-axis. D: Simulating evolution of the municipality population. The number of coins received between steps 80000 and 90000 was plotted in the y-axis and the number of coins at step 80000 was plotted in the x-axis.

$kN^{1.3-1.4}$, no plausible explanation could be given (3). In the present study, we examined a possible explanation that it could be related to the process of community formation by individuals.

The development of a local community was simulated by coin tossing under the condition that the probability of receiving a coin was proportional to the number of coins present in the box to the 1.3rd power. The simulation well-reproduced the actual situation (Fig. 4). The plot of the increment of number of coins against number of coins already present in the box fell on a straight line with a slope ~ 1.3 (Fig. 5B), which was close to the slope of the plot of number of inflow population against the number of residents, which was 1.23 for Japanese and 1.44 for non-Japanese (Fig. 5A). These slopes were close to the slope of the plot of the number of cases of syphilis or amebiasis against the population size (1.3–1.4) (Figs. 2C and 3D). If the pathogen carriers were randomly mixed within the inflow population, the equation $P = kN^{1.3-1.4}$ could be obtained, where P is number of the patients and N is the population size of the receiving community.

Conflict of interest None to declare.

REFERENCES

1. Yoshikura H. Impact of population size on incidence of rubella and measles in comparison with that of other infectious diseases. *Jpn J Infect Dis.* 2014; 67: 447-57.
2. Yoshikura H. Geo-demography of HIV/AIDS in Japan from 1985 to 2011 – Incidence and transmission mode under influence of population size/density. *Jpn J Infect Dis.* 2016; 69: 97-108.
3. Yoshikura H, Takeuchi F. Scale-free distribution of local infection cluster sizes of measles, rubella, syphilis and HIV/AIDS: correlation with size distribution of municipality population that was also scale-free. *Jpn J Infect Dis.* 2017; 70: 7-18.
4. Taya K. Revision of measles and rubella reporting to notification of all cases and second routine immunization with MR vaccine to the youth. *Infect Agents Surveillance Rep.* 2008; 29: 189-90. Japanese.
5. MHLW. Guidelines on the conduct of National Epidemiological Surveillance of Infectious Diseases. 1999. Available at <http://www.mhlw.go.jp/bunya/kenkou/kekkaku-kansen-shou11/dl/01_kansensho.pdf>. Retrieved in August 2016. Japanese.
6. National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division, Ministry of Health, Labour and Welfare, Japan. Measles, Japan, 2006 – 2007. *Infect Agents Surveillance Rep.* 2007; 28: 239'-40'.
7. Infectious Disease Surveillance Center, NIID: Available at <<http://www.mhlw.go.jp/shingi/2002/11/s1112-6c.html>>. Retrieved in August 2016. Japanese.
8. MHLW. Strategic Plan of Measles elimination. Available at <<http://www.mhlw.go.jp/shingi/2007/11/dl/s1102-5g.pdf>>. Retrieved in August 2016>. Japanese.
9. National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division, Ministry of Health, Labour and Welfare, Japan. Tuberculosis, as of 1998, Japan. *Infect Agents Surveillance Rep.* 1999; 20: 238'-9'.
10. National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division, Ministry of Health, Labour and Welfare, Japan. Shigellosis, Japan, 1993–1995. *Infect Agents Surveillance Rep.* 1996; 17:125'-6'.
11. Kobari I. Epidemiological and clinical observation of bacillary dysentery in the past three years – particularly on the *Shigella dysenteriae*. *J Jpn Assoc Infect Dis.* 1951; 23: 89-121. Japanese.
12. Yoshikura H. Strong correlation between annual incidence of male amebiasis and that of male homosexual HIV. *Jpn J Infect Dis.* 2016; 69:266-9.
13. National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division, Ministry of Health, Labour and Welfare, Japan. Enterohemorrhagic *Escherichia coli* infection as of April 2015. *Infect Agents Surveillance Rep.* 36:73'-4'.
14. Batty M. Cities and Complexity. Cambridge, MA: MIT Press; 2007. p. 1-16.
15. Ministry of Internal Affairs and Communications. Integrated administration of a large region, merger of cities, towns and villages. Available at <<http://www.soumu.go.jp/kouiki/kouiki.html>>. Japanese.