Original Article

Impact of Population Size on Incidence of Rubella and Measles in Comparison with That of Other Infectious Diseases

Hiroshi Yoshikura*
National Institute of Infectious Diseases, Tokyo 162-8640, Japan
(Received December 29, 2013. Accepted March 19, 2014)

SUMMARY: A strong dependency of rubella and measles epidemics on population size was confirmed by 2 types of plots: the cumulative frequency distribution of number of cases per prefecture and the slope of the log-log plots of number of cases per prefecture on the y-axis vs. prefecture population size on the x-axis. These parameters were found to be constant and unique to each infectious agent. The broader cumulative frequency distribution and steeper slope of the log-log plots were characteristic to measles and rubella, i.e., a higher population size was correlated with a disproportionate high incidence of measles and rubella. No such tendency was found in other infections with possible exceptions of pertussis and keratoconjunctivitis. The dependency of rubella and measles on population density was examined by log-log plots of patient number/population vs. population density, which revealed strong population density dependency of rubella; the dependency of measles on population density was equivocal.

INTRODUCTION

Around 1947, when Japan began to collect epidemiological data, it experienced large epidemics of measles every 2 years: 130,000–180,000 cases/year in peak years and 3,000–5,000 cases in years between. Since 1953, however, these peaks became less and less conspicuous, particularly with the start of routine measles vaccination in 1978 (1). However, Japan experienced an immense resurgence of measles infection in 2007, which affected many adults (2). Thereafter, measles elimination efforts in Japan were intensified in 2008 by adopting 2 doses of measles-rubella (M–R) vaccine in routine immunization and supplementary vaccination. The incidence of measles became barely observable after 1999 (4). The annual incidence decreased from 293 cases in 2008 to 87 cases in 2010. However, there has been notable resurgence with 2,328 cases in 2012 (5) and 14,269 cases (10,924 males and 3,345 females) as of week 46 in 2013 (http://www.nih.go.jp/niid/ja/diseases/ha/rubella.html [In Japanese]).

With regard to rubella, Japan experienced large epidemics every 5 years until 1992–1993, when the incidence began to decrease. Thereafter, the seasonal peak began barely observable after 1999 (4). The annual incidence decreased from 293 cases in 2008 to 87 cases in 2010. However, there has been notable resurgence with 2,328 cases in 2012 (5) and 14,269 cases (10,924 males and 3,345 females) as of week 46 in 2013 (http://www.nih.go.jp/niid/ja/diseases/ha/rubella.html [In Japanese]).

We examined the influence of the prefecture’s population size and density on the incidence of rubella, measles, and other infections.

MATERIALS AND METHODS

Graphic presentations: Graphic presentations possibly unique to this article will be explained below in order to avoid repetition of explanations.

Cumulative frequency distribution: Forty-seven prefectures in Japan were arranged in ascending order of a parameter, such as population size and number of cases, and ranked from 1 to 47, hereafter referred to as the ranking number. The number of cases per prefecture was plotted on the y-axis, and the prefecture’s ranking number was plotted on the x-axis. Thus, each prefecture was represented by a point with coordinates (number of cases, ranking number). If there were 2 or more prefectures reporting the same number of cases, they were represented by the prefecture with the highest ranking number so that the line connecting the points gives a cumulative frequency distribution curve. The detailed procedures are provided in the legend of Fig. 1 and in reference (6).

Patient number–population size plot: The patient number per prefecture was plotted on the y-axis, and the population per prefecture was plotted on the x-axis. An approximated correlation line was obtained by “power approximation” using the Excel file to obtain the slope of plots. If the patient number is proportional to population size, the slope will be 1 in tangent and 45° in angle. In the text, the slope is generally expressed in tangent only.

Patient incidence–population density plot: The number of cases per 1,000,000 population of the prefecture was plotted on the y-axis, and population per prefecture was plotted on the x-axis. An approximated correlation line was obtained by “power approximation” using the Excel file to obtain the slope of plots. If the patient number is proportional to population size, the slope will be 1 in tangent and 45° in angle. In the text, the slope is generally expressed in tangent only.

Correlation coefficient (CC): In the Excel file, a pair of columns was filled with 2 parameters, such as the number of cases per prefecture and population size, for
RESULTS

Cumulative frequency distribution of measles and rubella cases among prefectures: The cumulative frequency distribution of both measles and rubella cases per prefecture was sigmoid (Fig. 1, panels A and B). The frequency distribution made at log2 intervals (crosses connected with a dotted line in the both panels) resembled a log-normal distribution pattern. As measles and rubella epidemics declined from 2008 to 2012 and from 2008 to 2010, respectively, the curves shifted toward the left. For rubella, with the resurgence from 2011 to 2012, the curve shifted to the right. The width of the frequency distribution was estimated to be around 3 (see plots in closed squares for the measles epidemic in 2008 and plots in open circles for the rubella epidemic in 2012).

Cumulative frequency distribution of population size and population density among different prefectures in Japan: Since previous studies reported a correlation between measles epidemics and population size (7,8), we examined the cumulative frequency distribution of population size and density among the 47 prefectures in Japan (Fig 1, panels C and D, respectively). Both plots
showed log-normal-like distribution patterns with a width of the variation of around log10 for population size (panel C) and around 2 log10 for population density (panel D, closed circles). In panel D, the cumulative frequency distribution of population per habitable area is also shown. However, because the habitable areas of Japan account for only 33% of the total land mass, the percentage varied among prefectures. The range of variation was around log4 or 0.6 log10 for population per habitable area (http://www.japan-now.com/article/343318758.html [In Japanese]).

In panel E, population size is plotted on the x-axis against population density on the y-axis. The plots are distributed on a diagonal line with a slope of around 45°, suggesting that population density is approximately proportional to population size among the 47 prefectures in Japan (CC between population size and population density among prefectures was 0.89).

**Correlation between the number of measles or rubella cases and population size:** For both measles and rubella, the width of the cumulative frequency distribution of the number of cases per prefecture was about 3 fold wider than that of the population per prefecture (frequency distributions for measles in 2008 and rubella in 2012 are shown in Fig. 1 panels A and B, respectively, with population distribution in panel C). Plots of the numbers of measles and rubella cases vs. population size are shown in Fig. 2 (panels A and B, respectively). The slope of the approximated correlation line was 1.8–2.6 for measles and 2.6 for rubella (hereafter, the slope will only be expressed in tangent). When the patient number is proportional to population size, the slope of the approximation line will be 1 (as shown by the dotted line). Therefore, a large population size was associated with a
Fig. 3. Relation between patient number and population size for infectious diseases other than measles and rubella. For “sentinel-reporting” diseases, total number of patients reported from sentinels, “Table 10-2 Sentinel-reporting diseases (weekly); Number of cases per sentinel by week, prefecture, and sex (http://idsc.nih.go.jp/idwr/CDROM/Kako/H19/SyuList.html) Japanese.” are available. In this analysis, the former statistics were used because the latter becomes aberrant when incidence is very low. Consult figure legend for Fig. 2 for detail of graphics. The steepest slope was that of pertussis (1.27) followed by keratoconjunctivitis (1.14) (a thick line in each of panels A and B).

To determine if a steep slope was specifically characteristic to measles and rubella, other infectious diseases were similarly examined. Conducting such analysis was however not straightforward. While the reporting of measles and rubella cases has been mandatory since 2008 (i.e., physicians must report all cases), most other infectious disease with high incidences are “sentinel reporting diseases” under the Guidelines for National Epidemiological Surveillance of Infectious Diseases (http://www.mhlw.go.jp/bunya/kenkou/kekakusenshou11/dl/01_kansensho.pdf [In Japanese]). Available data for the sentinel reporting diseases are shown in Table 10-1: Sentinel-reporting diseases (weekly), while those for measles and rubella are shown in Table 1-1: Notifiable diseases: Number of cases by sex, prefecture, and week (http://www.nih.go.jp/niid/ja/all-surveillance/2270-idwr/nenpou/3359-syulist2011.html [In Japanese]).

To determine whether data obtained from the sentinel reporting system were suitable for comparisons with those obtained from the “all-the-case reporting” system, data regarding measles and rubella cases obtained before the switch of the reporting system in 2007 and those obtained after the switch in 2012 were compared. As shown in Fig. 2 (panels C and D), the approximated correlation lines before and after the switch (2007 and 2011, respectively) were almost parallel, which indicated that the data obtained from the “all-the-case” reporting system and those obtained from the sentinel reporting system are comparable, at least for comparison of the slopes. The approximated correlation lines for 14 other infections (i.e., gastroenteritis, erythema infectiosum, exanthema subitum, mumps, influenza, chickenpox, hand, foot, and mouth disease [HFMD], pertussis, pharyngoconjunctival fever, respiratory syncytial virus [RSV], group A streptococcal pharyngitis, herpangina, keratoconjunctivitis, and mycoplasma pneumonia) for the year 2011 (Fig. 3) had slopes around 1, with the exception of pertussis (slope = 1.27) and keratoconjunctivitis (slope = 1.14). The width of the cumulative frequency distribution curves of these infections was...
Rubella, Measles, Population Size Dependency

around log10 (Fig. 4, panel A), similar to that of the cumulative frequency distribution of population size (around log10). The curves of pertussis and keratoconjunctivitis were of an intermediate size, around 2 log10 (Fig. 4, panels A and B).

The slopes for RSV and mycoplasma pneumonia were slightly less steep than 1. Although the identities of all variables influencing these slopes remain unclear, a slope < 1 means that the higher the population size, the lesser is the patient number. Possible explanations for this phenomenon included disproportionate laboratory diagnoses among prefectures, inappropriate choice of sentinel points for these infections, uneven epidemics among prefectures during a particular year, and others. However, the main message of this paragraph is that no infections other than measles and rubella had slopes as steep as 2.

As observed in Fig. 4 (panels B and C) and Fig. 1 (panels A and B), the cumulative frequency curves horizontally shifted without changing the shape, i.e., the distribution pattern remained the same irrespective of the extent of the epidemic.

Correlation between the incidence of measles and rubella with population density: Plots of incidence number of patients per million population vs. population density for measles and rubella are shown in Fig. 5 (panels A and B, respectively). If the incidence is independent of population density, the plots will fall on a horizontal line, i.e., the slope will be zero. Here, data of measles infections in 2008 and rubella infections in 2012 were used for analysis because only in these years, the number of cases was sufficient to analyze the number of cases per population at the prefecture level.

With regard to rubella, the slope of the approximated correlation line was around 1, indicating that a high population density was associated with a high incidence of infection. For measles, plots were randomly distributed and the correlations were obscure. It should be noted here that the observed dependency of rubella on population density was observed in data from the large scale resurgence in 2012, while the measles data were derived from yearly epidemics. In addition, the observed absence of a correlation between measles patient number and population density could be attributable to the narrow variation of the habitable areas among prefectures (Fig. 1, panel D).
CCs between patient number and population size at the prefecture level: To confirm whether patient number and population size are correlated at the prefecture level, CCs, between patient number and population size among the 47 prefectures were calculated for each year, which are shown on the right side of graphs in panels A (measles) and B (rubella) of Fig. 2 (see figures that follow "CC"). One potentially interesting observation was that for measles, CC progressively increased from 0.67 in 2008 to 0.89 in 2010, in line with the steady decrease in the incidence of measles from 11,012 in 2008 to 293 in 2012. A possible explanation for this phenomenon is the advancement of immunization programs that gradually normalized immunity among prefectures and pushed the demographic factors in front. However, once the incidence of measles reached the recorded low (i.e., in 2012, 40 of 47 prefectures reported \(<\)10 cases), relying on CC became inappropriate because the incidence of measles at the prefecture level was too low and became stochastic (refer to Fig. 2 in reference [9]).

Regarding rubella, the highest CC value of 0.86 was attained during the resurgence in 2012, which may indicate that the resurgence evenly affected all the prefectures.

CCs between the number of cases and population size at the prefecture level: CCs between population density and patient number for measles were 0.59, 0.81, 0.80, 0.80, and 0.78 for the period 2008 to 2012, i.e., CC was at the most 0.8 for measles, while CCs were 0.83, 0.75, 0.86, and 0.67 for rubella for the period 2008 to 2011; however, the value increased to as high as 0.93 during the 2012 resurgence. These data indicate that the spread of rubella is more associated with a high population density than that of measles. The CC data are entirely consistent with the plot of measles incidence vs. population density shown in Fig. 5.

Stability of ranking of prefectures with regard to patient number over time: In Fig. 6, prefectures are ranked according to population size in ascending order from 1 to 47 and classified into 4 groups (see the column between columns for measles and those for rubella; see the inserted band in panel C in Fig. 1 as well). For each year, the prefectures are arranged in descending order of patient number. As observed, the prefectures with the greatest populations remained in the upper part of the table, while those with lower populations remained in the lower part. Hence, no large scale reshuffling of prefectures occurred.

Gender ratio: Fig. 7 examines the incidences of measles and rubella according to gender. In panels A and B, the number of male cases is plotted on the x-axis and the number of female cases is plotted on the y-axis, respectively. For a male-to-female ratio of 1:1, the plots will fall on a straight line with a slope of 1 crossing the origin (dotted line). As shown, the plots of measles were clustered around the straight line, while those of rubella were scattered below, indicating the predominance of male cases.

Difference between sexes with regard to the incidence of rubella is attributable to the past vaccination schedule in Japan, in which rubella vaccine was preferentially administered to girls, while routine immunization with M–R vaccine was only initiated in 2006 (5).
Fig. 6. Prefectures ranked for measles and rubella patient number from 2008 to 2012. Prefectures are ranked first according to the population size (column in the center); they are marked transparent, transparent with bold letters, moderately shaded and heavily shaded in ascending order (see also the bar-insert in panel C of Fig. 1). The prefectures are arranged according to the incidence of measles (on the left side) or rubella (on the right side) for each year.

Rubella, Measles, Population Size Dependency

Fig. 8 shows the age distribution of measles and rubella cases. For rubella (panel A), the distribution pattern remained almost identical from 2008 to 2010, with the highest incidence being among those aged <15 years, and it resembled that of measles (panel B). During the resurgence in 2011–2012 (diamonds and squares) (5), however, the incidence of rubella assumed an entirely different age distribution pattern. The age distribution curve of male cases was a large gently curved arch with a diffuse peak at 25–45 years, while that of the female cases was a curve with a sharp spike at 25–30 years. The sharp spike in the age distribution of females corresponds to the birth cohorts born within a few years before 1993 when the measles, mumps, and rubella (MMR) vaccine was suspended on account of being associated with aseptic meningitis (10). In addition, immunization regiments coverage of temporarily introduced individual immunization for male and female junior-high school students at hospitals in 1994–2003 was low (11). The incidence of rubella among cases aged >55 years was low even during the resurgence. Notably, this group represents birth cohorts born before 1960, when the incidence of measles in Japan first began to decrease (52,991 in 1966; 21,471 in 1981; and 6,716 in 1982). Although data from these periods are not available, this situation could be similar to that of rubella. Patient populations aged >50 years are considered to be immune to rubella and measles owing to
past epidemics; thus, the role of these groups in future epidemics will be insignificant despite their dominance in the Japanese population (see the age distribution of the Japanese population shown in crosses in panel A).

**DISCUSSION**

The correlation between population density and measles epidemic is well known (12–15). However, such dependency on population density was not evident for measles in this study. One main reason could be that the variation of population density of habitable areas is rather small in Japan. However, when different countries were compared, the influence of population density on measles was obvious (16,17). More recently, Finkenstäd et al. suggested a population density dependence of cyclicity with regard to measles epidemic (18). In a South African study, Sartorius et al. found an association of high population density with the risk of measles infection (19).

The observations made in the present report can be summarized as follows: (i) The slope of plots of patient number vs. population size and the width of the cumulative frequency distribution curves of patient number are parameters mutually related and invariant over time. (ii) The steep slope (around 2) of plots of patient number vs. population size and the wide cumulative frequency distribution pattern of patient number (around 3 log10) were unique to measles and rubella. For other infections so far examined, with the possible exceptions of pertussis and keratoconjunctivitis, the slopes were around 1 and the width of cumulative frequency distribution of the patient number was around log10.

These observations lead us to question of what determines population size dependency of measles and rubella, which was characterized by the steep slope of the plots of patient number vs. population size and the wide cumulative frequency distribution pattern of patient number. We reasoned that these 2 parameters are likely governed by modes of transmission of measles and rubella. According to a previous report (20), the principle mode of transmission for measles is “airborne by droplet spread; direct contact with nasal or throat secretions of infected persons,” that for rubella is “contact with nasopharyngeal secretions of infected persons through droplet spread or direct contact,” that for mumps is “droplet spread and by direct contact with saliva of an infected person,” that for pertussis is “direct contact with discharges from respiratory mucous membranes probably by droplets,” and that for influenza is “airborne spread among crowded populations in enclosed spaces predominates; transmission by direct contact through droplet spread.” Obviously, these modes of transportation are very similar in that all are mostly droplet-derived infections with direct contact favoring transmission. However, why only measles and rubella are strongly dependent on population size remains unclear. In these descriptions, there must be an important missing factor that differentiates measles/rubella from other infections.

One possible hint could be the observed slope of around 2 of the plots of patient number vs. population size, which was unique to measles and rubella. For N persons, the number of encounters will be proportional to the square of N. Therefore, if the transmission of a pathogen is strictly dependent on the encounter of 2 persons, the transmission event T will be proportional to the square of N, i.e., T = kN². If the equation is

![Fig. 7. Correlation between number of male patients and female patients in 47 prefectures. Panel A (rubella) and panel B (measles); horizontal axis: number of male patients per prefecture; vertical axis number of female patients per prefecture; dotted lines; lines with a slope of 45° crossing the origin of the coordinates. "cc" indicated after "year" is correlation coefficient between male and female patient numbers among prefectures. Panel C: cc between male patient numbers and female patient numbers is plotted in the y-axis against total patient number in the x-axis. Plots for rubella are in closed circles and those for measles in open triangles. Each plot corresponds to one of the years from 2008 to 2012. Note that the order of plots from the left to the right is not in the order of the year but in the ascending order of total number of patients. Number of patients in each year can be read from the value in the x-axis of the right most plots of the year in panel A (measles) and panel B (rubella) in Fig. 1.](image)
expressed in logarithmic values, it becomes $\log T = 2 \log N + \log k$, with a slope of 2, as observed for measles and rubella in Fig. 2. The transmission can be achieved more efficiently by routes other than direct contact, the number of encounters between the host and pathogen will be proportional to the population size $N$, and transmission event $T$ will be proportional to $E$, i.e., $T = kN$ or $\log T = \log N + \log k$, whose slope is 1 as observed in Fig. 3 for infections other than measles/rubella. Although the argument presented here is somewhat simplistic, it merits further exploration. Furthermore, whether such a hypothesis is compatible with the biological and ecological characteristics of these pathogens remains to be tested.

A further issue to be addressed is whether the population size dependency of measles and rubella interferes with the implementation of vaccination programs. If the incidence of measles or rubella is dependent on population size, the elimination of these viruses, originally estimated at less than 1 case per million population for measles (21), is daunting for countries with large populations. The World Health Organization (WHO) currently defines the elimination of measles as the interruption of endemic measles, which is theoretically correct (21); however, proving zero transmission is very costly, particularly for large, densely populated countries. In addition, measles and rubella are highly contagious and constantly imported and exported with the increased movement of humans due to globalization.

A schematic representation of the elimination process using the curve of the cumulative frequency distribution pattern of patient numbers per country and that of specific country populations is shown in Fig. 9. At the global level, the width of the cumulative frequency distribution of population ($W_p$) becomes enormous. For example, according to the statistics of the WHO Western Pacific Region, the population of China is 1,342,280,000; that of Japan is 127,450,000; that of Vietnam is 68,310,000; that of Thailand is 30,300,000; that of Singapore is 5,180,000; and that of Macao is 540,000. The difference in population size between China and Macao is about 2,500-fold; thus, the logarithmic value of the width of the cumulative frequency of population is more than 3-fold greater. Therefore, if the Japanese data presented in Fig. 1 is extrapolated to a global level, the width of cumulative frequency of patient number for measles will increase by
Fig. 9. Schematic representation of progression of control programs for measles, rubella, and other infections. The sigmoid curves represent cumulative distribution curves of population per country or number of patients per country. As control program progresses, the sigmoid curves move horizontally towards the left without changing their shape. The broader curve for measles and rubella allows less populated prefectures to eliminate the virus earlier (the left side downward slope) while leaving populated prefectures behind (the right upper side of the curve).

At least $3 \times 3$-fold on the logarithmic scale.

With the progress of elimination, the curve moves toward the left. To simplify matters, let us suppose that the curves move at a constant speed. The curve for other infections (dotted lines) will cross the elimination goal line in a shorter time because the width of the curve ($W-o$) is small. Meanwhile, for rubella and measles (solid lines), the front part of the curve will cross early; however, the tail of the curve lags behind because the width of the curve ($W-m/r$) is large; therefore, it will take a longer period of time from the first success of disease elimination in a less populated country to the completion of elimination among highly populated countries. The current global situation fits this picture exactly; while measles elimination occurs relatively rapidly in regions of the United States, where countries with high population densities and highly populated countries do not exist, it is difficult in other regions with high population densities and high population size (17). It is, however, important to note that infections represented by curves with narrower widths are not necessarily eliminated easily because they may be rather immobile despite any intervention, such as that observed in influenza epidemics.

The above analysis does not necessarily imply that the elimination of measles and rubella elimination is impossible, because the dependency on population size means that once the size of the susceptible population is reduced to a certain level, the elimination of both measles and rubella can be accelerated. Osaka, which is one of the most populated prefectures in Japan, had an incidence of measles of 4 cases per 8,865,000 population (0.45/million) in 2012 (3). I entirely agree with Rozhnova et al. (22) in that "for rubella, reducing $R_0$ by vaccinating or declining birth rate unambiguously results in higher extinction probabilities."

Acknowledgments The author’s thanks are due to Drs. K. Taya, K. Kinoshita, and K. Yamashita, Infectious Disease Surveillance Center, National Institute of Infectious Diseases for information of data source.

Conflict of interest None to declare.

REFERENCES