Laboratory and Epidemiology Communications

Infectious Disease, Population Size/Density, and Tax Income of Prefectures

Hiroshi Yoshikura*
National Institute of Infectious Diseases, Tokyo 162-8640, Japan

Communicated by Ichiro Kurane

This communication will show how infectious disease is influenced by human demography and economic activity. AIDS, rubella, measles, and syphilis were positively correlated with population size/density, while scrub typhus, legionellosis, and enterohemorrhagic *Escherichia coli* (EHEC) infections were negatively correlated.

In Japan, under the Law Concerning the Prevention of Infections and Medical Care for Patients of Infections, physicians who have diagnosed infectious diseases in category I–IV and some in category V shall report all the cases without delay (http://law.e-gov.go.jp/htmldata/H10/H10HO114.html). Such infections are conveniently called "all the case notifiable infections." In the present analysis, "all the case notifiable infections" reported in relatively large numbers (>100 cases/year) were selected.

Fig. 1 plots the incidence (number of patients per 1,000) in the horizontal axis (x-axis) and ranking number of prefectures starting from the prefecture with the lowest incidence in the vertical axis (y-axis). The plot is thus a cumulative frequency distribution of disease incidence per prefecture. EHEC, shigellosis, and tuberculo-
sis showed near normal distribution, while all others (i.e., measles, rubella, syphilis, AIDS, severe invasive streptococcal infections [TSLS], amebiasis, hepatitis A [HAV], hepatitis E [HEV], legionellosis, and scrub typhus [Tsutsugamushi disease]) showed near log-normal distribution. The downward arrowheads above the graphs indicate the approximate position of the arithmetic mean of the incidences.

Fig. 2 shows what are called "incidence-population size/density" plots. Circles represent prefectures. The vertical axis is the population density (×100/sq. km), and the horizontal axis is the population size (×100,000). Prefectures with large population sizes are on the right side, and those with high population density are on the upper side of the graph. Tokyo, Osaka, and Kanagawa, highly populated prefectures, are clustered in the upper-right, and Aomori, Akita, and Okinawa, lowly populated prefectures, are in the lower left. The disease incidence (number of patients per 1,000) of prefectures is represented by the circle sizes. The prefectures whose incidence is higher than the arithmetic mean among prefectures are represented by large circles (prefectures on the right side of the arrowheads in Fig. 1), and those whose incidence is lower than the arithmetic mean are represented by small circles (prefectures on the left side of the arrowheads in Fig. 1). For AIDS, measles, syphilis, rubella, and amebiasis, large circles are clustered in the upper right, while those for legionellosis, scrub typhus, and EHEC (Fig. 2) are clustered in the lower left. The clustering is not obvious for tuberculosis, TSLS, HAV, HEV, and shigellosis.

For all the infections (including HAV and HEV, whose 2012 data are not shown), the plot pattern is similar between 2012 and 2013, indicating relative pattern stability at least in a short time span.

How could the population size/density have influenced the geographical distribution of infections? It may have been human activities linked to human demography rather than the population size/density itself that directly influenced the infectious disease epidemiology.

Prefectures' tax income is a measure of their economic activities. Fig. 3A shows the relationship between the tax income per person in 2009 and the population size density. Prefectures whose tax income was larger than the arithmetic mean (¥37,600) are represented by large circles, and those with smaller tax income are represented by small circles. Large circles are clustered in the upper right, and small circles are clustered in the lower left, showing that tax income per person is higher for more populated prefectures and lower for less populated prefectures. Fig. 3B-1 and 3B-2 show the relationship between tax income and population size and that between tax income and population density, respectively. The plots are rightward ascending; the correlation coefficient (cc) between tax income and population size is 0.823 and the cc between tax income and population density is 0.782; the higher the population size/density, the higher the prefectures' tax income. In Fig. 3, the tax income per prefecture is plotted in the horizontal axis, and the disease incidence is plotted in the vertical axis. The plot distribution is rightward ascending for AIDS and measles (Figs. 3C-1 and 3C-2), and the 2 parameters are mutually correlated with cc's of 0.737 for AIDS and 0.691 for measles. The plot distribution for tuberculosis is almost horizontal (Fig. 3C-3), and there is no correlation between tuberculosis incidence and tax income (cc: 0.260).

Fig. 4A compares age distribution of prefectures. Percentages occupied by age groups 0–4 years, 15–34 years, 35–65 years, and >65 years were calculated for individual prefectures. Within each age group, prefectures were divided into those whose percentage was higher than the arithmetic mean of each age group (represented by large circles in the figure) and those whose percentage was lower than the arithmetic mean (represented by small circles). In age groups 15–34 and

*Corresponding author: Mailing address: Policy Planning and Communication Division, 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

---

Accepted May 22, 2015.
DOI: 10.7883/yoken.JJID.2015.110

---

446
Fig. 1. Cumulative frequency distribution of various infections. The data were derived from the 2013 Annual Report of the National Epidemiological Surveillance of Infectious Diseases (http://www.nih.go.jp/niid/ja/survei/2270-idwr/nenpou/5281-syulist2013.html). Infections used for the analysis were selected from those requiring notification of all the cases that physicians diagnosed and those whose annual number of cases exceeds 100. Arrows indicate the approximate position of the arithmetic mean of incidence (number patients per 1,000); Measles: 0.00077; Rubella: 0.066; Syphilis: 0.0063; AIDS: 0.0079; Severe Invasive Streptococcal Infections (TSLS): 0.0071; Amebiasis: 0.0066; HAV: 0.00082; HEV: 0.00074; Legionellosis: 0.0097; Scrub typhus: 0.0043; EHEC: 0.0035; Shigellosis: 0.201; Tuberculosis: 0.196. All the infections followed near log-normal distribution except EHEC, shigellosis, and tuberculosis whose distribution was more compatible with normal distribution.

35–64 years, large circles are clustered in the upper right, while in the age group >65 years, large circles are clustered in the lower right. These indicate that the working age population, 15–34 and 35–64 years, is concentrated in populated urban prefectures, and the aging population is concentrated in less populated rural prefectures. In the age group 0–14 years, large circles and small circles are intermingled; however, circles for Tokyo, Osaka, Kanagawa, and Saitama that are clustered in the upper right are small, indicating that birth rate in the most populated prefectures in Japan is low.

Now, the incidence-population size/density plots in Fig. 2 can be examined for compatibility with the age distribution in Fig. 4A. (i) AIDS, syphilis, and amebiasis: Clustering of large circles in the upper right is compatible with the fact that these infections, including amebiasis, are linked to sexual activities of middle-aged and young people. (http://www.nih.go.jp/niid/images/iasr/35/415/graph/f4154j.gif; http://www.nih.go.jp/niid/ja/id/742-disease-based/ha/syphilis/idsr-topic/5404-tpc420-j.html; http://idsr.nih.go.jp/iasr/28/326/tpc326-j.html). (ii) Measles and rubella: Clustering of large circles in the upper right is compatible with the fact that the peak age of rubella patients is currently 30–40 years (http://www.nih.go.jp/niid/images/iasr/34/398/graph/f3983j.gif) and that measles is transmitted to the susceptible age group <2 years by adults (1) (http://idsr.nih.go.jp/iasr/31/360/graph/f3603j.gif). (iii) Legionellosis and scrub typhus: Clustering of large circles in the lower left is compatible with the fact that most patients are aged 50–90 years with the peak in 60 years for legionellosis (http://www.nih.go.jp/niid/ja/id/687-disease-based/ra/legionella/idsr-topic/3611-tpc400-j.html) and 70–74 years for scrub typhus. Scrub typhus patients are mostly farmers (http://idsr.nih.go.jp/iasr/31/363/tpc363-j.html). (iv) EHEC and shigellosis: Though they are similar as pathogens, while EHEC’s large circles are clustered in the lower left, shigellosis’s large circles are clustered in the upper right.
Fig. 2. Incidence of infections as a function of population size and population density. Each symbol represents a prefecture, whose coordinates are determined by the population size in the horizontal axis (× 100,000) and population density in the vertical axis (× 100/sq. km). Incidence (number of patients per 1,000) of the infections is represented by the size of the symbols. Larger symbols represent prefectures whose incidence was higher than the arithmetic mean of the patient number per prefectures, and smaller symbols prefectures with incidence lower than the arithmetic mean. The approximate position of the mean of incidences among prefectures is shown in Fig. 1 with downward arrowheads on the top of each graph (See Fig. 1). Correlation coefficients between the incidence and population size and those between the incidence and population density, which were calculated by using the 2013 data, were respectively 0.794 and 0.702 for measles, 0.610 and 0.704 for rubella, 0.612 and 0.715 for syphilis, 0.735 and 0.812 for AIDS, 0.175 and 0.184 for severe invasive streptococcal infections (TSLS), 0.527 and 0.523 for amebiasis, 0.210 and 0.193 for HAV, 0.225 and 0.083 for HEV, −0.148 and −0.144 for Legionella, −0.266 and −0.239 for scrub typhus (Tsutsugamushi disease), −0.202 and −0.120 for Enterohemorrhagic Escherichia Coli (EHEC), 0.346 and 0.625 for Shigellosis, and 0.289 and 0.33 for tuberculosis. The plots for 2012 and those 2013 are arranged side by side except HAV and HEV, for which only the plots for 2013 are shown due to space constraints.
Fig. 3. Relation between prefectures’ tax income and population size/density (A, B-1, and B-2) and relation between the tax income and incidence of AIDS (C-1), measles (C-2) and tuberculosis (C-3). A: Tax per person and population size/density. The prefectures’ tax income data (2009) were derived from http://www.soumu.go.jp/main_sosiki/jichi_zeisei/czaisei/czaisei_seido/pdf/ichiran06_h20_21.pdf. For the plot, individual inhabitant tax and corporation tax were added. Vertical axis: population density \((\times 100/\text{sq. km})\), horizontal axis: population size \((\times 100,000)\), larger symbols: prefectures with tax income larger than the mean of all the prefectures’ tax income (mean ¥37,600 per person; range: ¥24,100–¥83,200). Correlation coefficient between tax income and population size was 0.823 and that between tax income and population density was 0.791. B-1: Tax per person \((\times 10,000\text{¥})\) in Y axis vs. population size \((\times 100,000)\) in X-axis; correlation coefficient: 0.823. Lines were drawn by using the power approximation on Microsoft Excel 2010. B-2: Tax per person \((\times 10,000\text{¥})\) in Y axis vs. population density \((\times 100/\text{sq. km})\) in X-axis; correlation coefficient: 0.782. C: Tax per person \((\times 10,000\text{¥})\) in X axis vs. incidence of AIDS (C-1), measles (C-2), or tuberculosis (C-3) (per 100,000) Correlation coefficient for tax and AIDS was 0.737, that for tax and measles 0.691, and that for tax and tuberculosis 0.260. Incidence of AIDS and that of measles were significantly correlated with the tax income, while incidence of tuberculosis and the tax was not. The lines were drawn using the linear approximation on Excel file 2010.

The difference could be attributable to the fact that shigellosis was largely borne by the young generation who went abroad (20–40 years) (http://idsc.nih.go.jp/iasr/30/358/tpc358-j.html), while for EHEC, the proportion of symptomatic infection was high among age groups <10 years and >60 years (http://www.nih.go.jp/niid/ja/ehec/552-idsc/iasr-topic/4622-tpc411-j.html). (v) H1N1 2009 pdm in 2009 (Fig. 4B): This was the only 1 influenza infection in the category of “all the case notifiable infections.” Large circles are clustered in the lower left, which is a pattern of infections in older adults (compare H1N1 2009 pdm plot in Fig. 4B-1 with the >65 year plot in Fig. 4A). However, the patients’ age was clustered in 5–24 years, as shown in Fig. 4B-2. Close examination, however, reveals resemblance between the H1N1 2009 pdm plot and the 0–14 year plot in Fig. 4A. The overall pattern of the H1N1 2009 pdm plot was probably represented by the age group younger than 14 years (actually, as shown in Fig. 4B-2, patients aged above 15 years were a minority). This example warns against use of the incidence-population size/density plots in isolation.

Analysis presented in this paper illustrates how population size/density, tax income, and age distribution are related with infectious disease incidences, though they may not be in a direct causal relationship. Control/elimination of measles, rubella, AIDS, syphilis, amebiasis, etc., which are under strong influence of human activity, will be difficult, though not impossible, unless targeted measures are taken. Consideration of geo-demographic factors will be indispensable for such enterprises.

Meanwhile, the present analysis confirmed problems associated with the aging population advocated by Masuda (2). In Japan, rural prefectures are aging and their tax income is decreasing, while they have to confront infections that are relatively rare in populated prefectures.

Geo-demography will be an important consideration for policy making on infectious diseases because the surrounding situations are uneven in Japan, and the unevenness may be aggravated in future.
Fig. 4. Age distribution among 47 prefectures. A: The population is divided into 4 age groups, 0–14, 15–34, 35–64, and >65 years. For each prefecture, percentage of the population belonging to each age group is calculated. Prefectures whose percentage was larger than the average are represented by large circles and those whose percentage was smaller are represented by small circles. Each circle represents a prefecture, whose coordinates are determined by the population size in the horizontal axis ($\times 100,000$) and population density in the vertical axis ($\times 100$ per sq. km). The mean and range for individual age groups were respectively 13.2% and 11.1%–17.6% for <14 years; 20.2% and 16.9%–24.7% for 15–34 years; 40.9% and 39.3–42.9% for 35–65 years; and 25.5%, and 17.1%–30.7% for >65 years. Data source for the age distribution among prefectures was Table 4: Population by Sex and Sex ratio for prefectures—total population, Japanese population, October 1 2012 (http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001118081). B-1: incidence-population size/density plot of H1N1 2009 pdm influenza. B-2: Age distribution of the H1N1 2009 pdm influenza patients. Vertical axis: number of patients. Horizontal axis: age groups.

Conflict of interest None to declare.

REFERENCES