

# Predicting the characteristics of the aetiological agent for Kawasaki disease from other paediatric infectious diseases in Japan

Y. NAGAO<sup>1</sup>\*, C. URABE<sup>2</sup>, H. NAKAMURA<sup>3</sup> AND N. HATANO<sup>2</sup>

<sup>1</sup> *Japan Community Health Care Organization, Osaka Hospital, Fukushima, Osaka, Japan*

<sup>2</sup> *Institute of Industrial Science, the University of Tokyo, Komaba, Meguro, Tokyo, Japan*

<sup>3</sup> *National Institute for Fusion Science, Oroshi-cho, Toki, Gifu, Japan*

*Received 25 December 2014; Final revision 17 April 2015; Accepted 13 May 2015;  
first published online 23 July 2015*

## SUMMARY

Although Kawasaki disease (KD), which was first reported in the 1960s, is assumed to be infectious, its aetiological agent(s) remains unknown. We compared the geographical distribution of the force of infection and the super-annual periodicity of KD and seven other paediatric infectious diseases in Japan. The geographical distribution of the force of infection, which was estimated as the inverse of the mean patient age, was similar in KD and other paediatric viral infections. This similarity was due to the fact that the force of infection was determined largely by the total fertility rate. This finding suggests that KD shares a transmission route, i.e. sibling-to-sibling infection, with other paediatric infections. The super-annual periodicity, which is positively associated with the sum of an infectious disease's incubation period and infectious period, was much longer for KD and exanthema subitum than other paediatric infectious diseases. The virus for exanthema subitum is known to persist across the host's lifespan, which suggests that the aetiological agent for KD may also be capable of persistent infection. Taken together, these findings suggest that the aetiological agent for KD is transmitted through close contact and persists asymptotically in most hosts.

**Key words:** Emerging infections, epidemiology, infectious disease epidemiology, Kawasaki disease, mathematical modelling.

## INTRODUCTION

Kawasaki disease (KD), which is the most common cause of paediatric vasculitis in developed countries, was first reported in Japan in the 1960s [1]. Following three large epidemics of KD, which occurred in 1979, 1982 and 1986, the incidence of KD has been steadily increasing in Japan [2]. Although the aetiology of KD has not been identified, KD is

assumed to be infectious based on the presence of seasonality [3] and spatio-temporal clustering during epidemics [4]. In addition, KD shares many clinical manifestations (e.g. fever, conjunctivitis, rash, cervical adenopathy) with other viral infections [5]. The rarity of KD in infants aged <6 months indicates the presence of an effective passive immunity [6]. The recurrence of KD is rare [7], which suggests that KD may be caused by a limited number of – if not a single – aetiological agents. In recent years, a number of large-scale epidemiological studies have enhanced our understanding of the genetic susceptibility to KD [8–10]. By contrast, however, the aetiological agent(s) of KD has not been identified.

\* Author for correspondence: Dr Y. Nagao, Japan Community Health Care Organization, Osaka Hospital, 4-2-78 Fukushima, Fukushima, Osaka, 553-0003 Japan.  
(Email: in\_the\_pacific214@yahoo.co.jp)

Table 1. Kawasaki disease and paediatric infectious diseases compared in the present study and their mean patient ages (years) in Japan measured between 2000 and 2010

Disease	Main aetiological agents	Vaccine	National crude mean age	National adjusted mean age	Range of prefectural crude mean age	Range of prefectural adjusted mean age
Kawasaki disease	Unknown	Not available	2.6	2.4	2.1–2.7	2.1–2.6
Exanthema subitum*	Human herpes virus 6 & 7	Not available	0.91	0.91	0.78–1.1	0.78–1.1
Herpangina	Enteroviruses	Not available	3.4	3.4	2.3–4.3	2.4–4.2
Hand, foot and mouth disease	Enteroviruses	Not available	3.7	3.6	2.5–4.1	2.5–4.1
Chickenpox	Varicella-zoster virus	Voluntary	3.7	3.7	3.1–4.3	3.0–4.3
Pharyngoconjunctival fever	Adenoviruses	Not available	4.6	4.5	3.5–5.7	3.4–5.7
Erythema infectiosum	Human parvovirus B19	Not available	6.1	6.1	5.5–6.7	5.4–6.7
Group A streptococcus	<i>Streptococcus pyogenes</i>	Not available	7.1	7.1	6.1–8.4	6.1–8.3

\* Synonym for roseola infantum.

KD can cause a coronary artery lesion and a subsequent fatal myocardial infarction in 1–2% of patients [11, 12], unless treated promptly with high-dose intravenous immunoglobulin (IVIG) [13]; however, 40% of Japanese KD patients who are treated with IVIG therapy still develop transient coronary artery lesions [14]. IVIG therapy is not only expensive but also harbours the potential for several adverse effects. The identification of the aetiological agent for KD is therefore critical to the development of less expensive and more specific therapies. In the present study, we compared KD with other paediatric infectious diseases with known aetiologies in order to characterize the likely aetiology of KD. A comparison of the epidemiological profiles of multiple diseases can shed light on the determinants of transmission of the individual diseases [15]. Here, we focused on two epidemiological characteristics, the geographical distribution of the force of infection (FOI) and the periodicity, for KD and each of the seven other paediatric infections.

## METHODS

### KD data

We obtained the KD data from the Nationwide Surveillance of Kawasaki Disease (NSKD), which includes almost all of the reported cases of KD stratified by age across all 47 prefectures of Japan [16] since 1979 [17].

### Paediatric infectious disease data

The data for nine infectious diseases stratified by age have been reported weekly to the National Epidemiological Surveillance of Infectious Diseases (NESID) from approximately 3000 sentinel paediatric clinics since 1999 (e.g. 2978 clinics in 2000 and 3028 clinics in 2010) [18]. Our analysis excluded infectious gastroenteritis, which is caused by multiple heterogeneous pathogens, and mumps, for which the information about vaccine coverage cannot be obtained. Therefore, we analysed the data for seven paediatric infectious diseases (Table 1). Because the allocation of sentinel clinics is not precisely proportional to the population size, the absolute incidence of disease cannot be estimated from these data. Instead, we estimated the mean patient age from the patient age distribution (see Supplementary Fig. S1).

### Estimation of the mean patient age

The mean patient age of each infectious disease, including KD, varies by season [19]. We pooled the

data reported between 2000 and 2010 from both the NSKD and the NESID to circumvent any interference from this seasonal variation. We estimated the crude mean patient age for KD and the seven other paediatric infectious diseases using the following equation [17]:

$$cA_{ij} = \frac{\sum_{k=1}^{11} \sum_{m=2000}^{2010} (M_k \times N_{ijkm})}{\sum_{k=1}^{11} \sum_{m=2000}^{2010} (N_{ijkm})}, \quad (1)$$

where  $cA_{ij}$  denotes the crude mean age for the  $i$ th disease in the  $j$ th prefecture,  $M_k$  represents the midpoint age for the  $k$ th age group, and  $N_{ijkm}$  represents the number of patients with the  $i$ th disease from the  $k$ th age group in the  $j$ th prefecture in year  $m$ . The definitions of the 11 age groups (Supplementary Fig. S1) were consistent across the diseases and throughout the study period.

To eliminate interference due to the demographic structure of the data, the mean patient age for each disease was adjusted [17] using the following equation:

$$aA_{ij} = \frac{\sum_{k=1}^{11} \sum_{m=2000}^{2010} (M_k \times N_{ijkm} \times S_k / C_{jkm})}{\sum_{k=1}^{11} \sum_{m=2000}^{2010} (N_{ijkm} \times S_k / C_{jkm})}, \quad (2)$$

where  $aA_{ij}$  represents the adjusted mean patient age for the  $i$ th disease in the  $j$ th prefecture,  $S_k$  represents the proportion of the population in the  $k$ th age group in the standard population structure (i.e. Japan's population structure in 2000), and  $C_{jkm}$  represents the proportion of the population in the  $k$ th age group in the  $j$ th prefecture in year  $m$ .

In addition, we estimated the national-level mean patient ages using the following equations:

$$\text{national\_}cA_i = \frac{\sum_{j=1}^{47} \sum_{k=1}^{11} \sum_{m=2000}^{2010} (M_k \times N_{ijkm})}{\sum_{j=1}^{47} \sum_{k=1}^{11} \sum_{m=2000}^{2010} (N_{ijkm})}, \quad (3)$$

$$\text{national\_}aA_i = \frac{\sum_{j=1}^{47} \sum_{k=1}^{11} \sum_{m=2000}^{2010} (M_k \times N_{ijkm} \times S_k / C_{jkm})}{\sum_{j=1}^{47} \sum_{k=1}^{11} \sum_{m=2000}^{2010} (N_{ijkm} \times S_k / C_{jkm})}, \quad (4)$$

where  $\text{national\_}cA_i$  and  $\text{national\_}aA_i$  represent the crude and adjusted mean patient ages of the  $i$ th disease at the national level, respectively.

## Estimation of the FOI

The distribution of the FOI often identifies the risk factors for transmission [20] and may also reveal the mode of transmission. In the present study, a surrogate measure of the FOI was estimated as the inverse of the mean patient age as shown in the following equation:

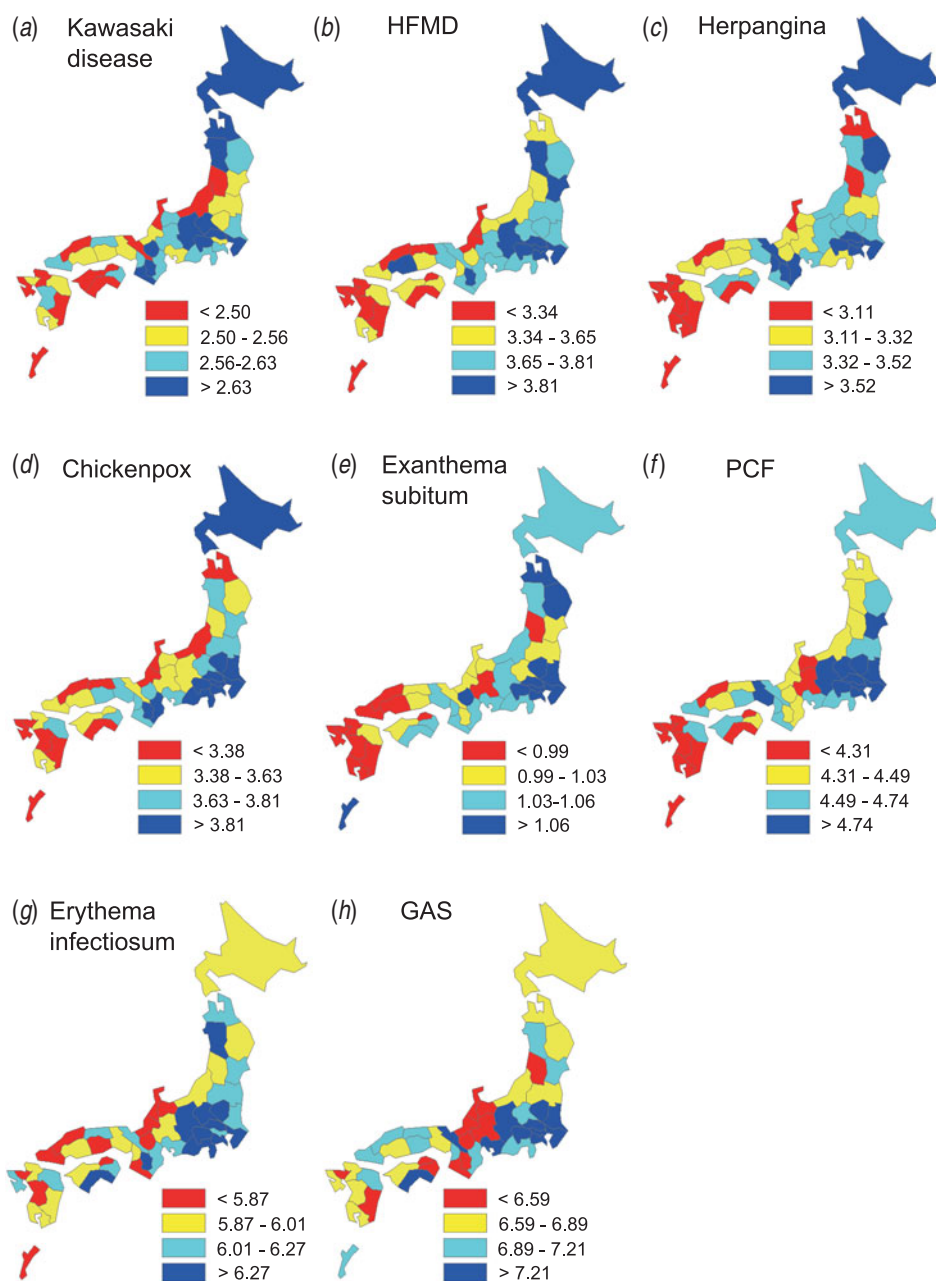
$$A = 1/\lambda, \quad (5)$$

where  $A$  represents the mean patient age and  $\lambda$  represents the FOI for an infectious agent that manifests as an acute illness and confers lifelong immunity [21]. The FOI estimated by equation (5) has been shown to be a good approximation of more elaborate mathematical techniques used to compute the FOI [22]. The FOI and the mean patient age have been shown to be negatively correlated even in a disease caused by multiple viral strains [23].

## Geographical analysis

The geographical distribution of the mean patient age was similar between KD and the other paediatric infectious diseases (Fig. 1). To identify the determinants of this shared geographical distribution, we estimated the correlation between the adjusted mean patient ages and climatic/socioeconomic variables. The following climate data was obtained from the University Corporation of Atmospheric Research [24] as described previously [17]: mean temperature ( $^{\circ}\text{C}$ ), precipitation (mm/day), average vapour pressure (mmHg), and the average vapour pressure deficit (which represents aridity, mmHg). Thirty-seven socioeconomic variables were obtained from the Ministry of Internal Affairs and Communications of Japan [25], including seven demographic variables, eight education variables, eight health variables, seven infrastructure variables, four variables related to the standard of living, and three landscape variables (Supplementary Table S1). The values of these variables, which were surveyed at least three times between 2000 and 2010, were averaged for each prefecture. Mitsubishi Tanabe Pharma (Osaka, Japan), the sole retailer of varicella vaccine in Japan, provided the uptake rate of the varicella vaccine at the prefecture level for each year between 2000 and 2010.

A rank correlation analysis was used to screen the climatic/socioeconomic variables that exhibited a statistically significant ( $P < 0.05$ ) rank correlation with the adjusted mean patient age of all the diseases. We



**Fig. 1.** Similarity in the geographical distribution of the crude mean patient age for Kawasaki disease and seven paediatric infectious diseases. The prefectures are categorized from red to blue in ascending order of the crude mean patient age. The distributions of the adjusted mean patient ages were similar to the results presented here. Only data from the main part of each prefecture is presented; minor islands were omitted in this and subsequent figures. Okinawa, which is 650 km away from the main island, is shown close to the main island in this representation. HFMD, Hand, foot and mouth disease; PCF, pharyngoconjunctival fever; GAS, group A streptococcus.

applied multivariate regression analyses to these pre-selected variables. The adjusted mean patient age was regressed against the explanatory variables using a conventional multivariate linear regression analysis. Then, to account for spatial autocorrelation, we applied a spatial multivariate regression analysis [26]

using the following equation:

$$Y = \rho UY + X\beta + \text{const.}, \quad (6)$$

where  $Y$  represents the dependent variable vector,  $X$  represents the independent variable matrix,  $\rho$  represents the spatial autoregressive parameter,  $\beta$  denotes

the regression coefficient vector, and  $U$  represents the spatial adjacency matrix. Each element of  $U$  was assigned a value of 1 if the two prefectures were geographically adjacent or connected by a bridge or tunnel, and 0 if the two prefectures were not adjacent or connected. The use of the crude mean patient age ( $cA_{ij}$ ) instead of the adjusted mean patient age ( $aA_{ij}$ ) did not result in qualitatively different results (data not shown). We used Stata v. 11.1 (StataCorp., USA) for all statistical analyses.

### Estimation of the periodicity

We also compared the periodicity between KD and other paediatric infectious diseases. The super-annual periodicity ( $T$ ) in the number of cases was determined by the following equation:

$$T \cong 2\pi[(D + D')A]^{\frac{1}{2}}, \quad (7)$$

where  $D$  and  $D'$  represent the incubation and infectious periods, respectively, and  $A$  represents the mean patient age of the infectious disease [21]. This relationship holds for persistent infections (e.g. *Plasmodium falciparum*) as well as diseases caused by complex interactions between multiple strains (e.g. dengue haemorrhagic fever) [27]. We tested whether this equation applies to paediatric infectious diseases, and then applied this equation to predict the nature of the aetiological agent for KD.

### Time-series analysis

The monthly number of cases of KD in Japan were available from the NSKD. We converted the weekly number of the seven paediatric infectious disease cases, which were reported from the sentinel paediatric clinics to the NESID between 2000 and 2010, into the monthly number of cases. The monthly number of KD cases and the seven other paediatric diseases were transformed as the proportion to the maximum number of cases recorded during the study period, and were subsequently analysed by the wavelet method [28] using the Morlet wavelet [29]; we followed the standard wavelet procedure used in many recent epidemiological studies (e.g. [23]). We used Mathematica 9.0 (Wolfram, USA) for this analysis.

### Estimation of the probability of KD manifestation

We expressed the proportion of individuals who are naive to the aetiological agent for KD as  $V$ . The

rate of change of  $V$  at the age of  $t$  years can be expressed as:

$$dV = -\lambda V dt, \quad (8)$$

where  $\lambda$  represents the FOI. Equation (5) holds for a persistent infection if the incubation period is not very long. Therefore, we solved equation (8) using equation (5) and a condition of  $V(0) = 1$  into the following equation:

$$V = \exp(-\lambda t) = \exp(-t/A), \quad (9)$$

where  $A$  represents the mean patient age. Based on this equation and epidemiological and demographic data from Japan, we estimated the probability of developing KD when infected with its aetiological agent.

## RESULTS

### Geographical analysis

The distribution of the mean patient age was similar between KD and the other paediatric infectious diseases (Fig. 1); there was a statistically significant rank correlation between all pairs of these diseases with respect to the prefectural mean patient age (Supplementary Table S2). The rank correlations between the adjusted mean patient ages and the climatic/socioeconomic variables were estimated (Supplementary Table S3). This screening process revealed eight variables that were significantly correlated with the adjusted mean patient ages of all the diseases (Table 2). Four of these variables were related to healthcare, and hence were mutually correlated. We excluded non-unique variables to avoid multi-collinearity effects (see uniqueness in Table 2), and included the varicella vaccine uptake rate as a covariate. The final statistical model for each disease was determined by a step-wise addition and elimination of these variables (Table 3).

The health insurance paid per insuree remained as a significant negative contributor to the adjusted mean patient age for most of the diseases. The total fertility rate (TFR) remained as a significant negative contributor to the adjusted mean patient age for all the diseases except group A streptococcus (GAS). These results suggest that the availability/utilization of health services, and more importantly, the TFR (Fig. 2) are critical factors in the transmission of KD and paediatric infectious diseases. Notably, Okinawa prefecture, which has the largest TFR and lowest health insurance paid per insuree, emerged as an outlier for some diseases (Supplementary Figs S2, S3).

Table 2. Variables which exhibited statistically significant rank correlations with the adjusted mean patient age of Kawasaki disease and paediatric infectious diseases ( $n = 47$ )

Variable (unit)	Uniqueness*	Kawasaki disease	Exanthema subitum	Herpangina	HFMD	Chickenpox	PCF	Erythema infectiosum	GAS
Total fertility rate (dimensionless)	0.5141	-0.46 ( $P = 0.0011$ )	-0.46 ( $P = 0.0010$ )	-0.67 ( $P < 0.0001$ )	-0.63 ( $P < 0.0001$ )	-0.60 ( $P < 0.0001$ )	-0.47 ( $P = 0.0008$ )	-0.44 ( $P = 0.0021$ )	-0.33 ( $P = 0.0237$ )
Pupil per class (persons)	0.1920	0.44 ( $P = 0.0021$ )	0.31 ( $P = 0.0365$ )	0.33 ( $P = 0.0224$ )	0.40 ( $P = 0.0049$ )	0.46 ( $P = 0.0013$ )	0.39 ( $P = 0.0074$ )	0.30 ( $P = 0.0440$ )	0.33 ( $P = 0.0236$ )
Hospital (/100 000)	0.1318	-0.45 ( $P = 0.0015$ )	-0.50 ( $P = 0.0004$ )	-0.58 ( $P < 0.0001$ )	-0.64 ( $P < 0.0001$ )	-0.56 ( $P < 0.0001$ )	-0.55 ( $P = 0.0001$ )	-0.48 ( $P = 0.0007$ )	-0.42 ( $P = 0.0032$ )
Bed (/100 000)	0.1038	-0.42 ( $P = 0.0036$ )	-0.56 ( $P < 0.0001$ )	-0.61 ( $P < 0.0001$ )	-0.68 ( $P < 0.0001$ )	-0.64 ( $P < 0.0001$ )	-0.64 ( $P < 0.0001$ )	-0.53 ( $P = 0.0001$ )	-0.44 ( $P = 0.0022$ )
Nurse (/100 000)	0.0173	-0.49 ( $P = 0.0004$ )	-0.64 ( $P < 0.0001$ )	-0.72 ( $P < 0.0001$ )	-0.75 ( $P < 0.0001$ )	-0.72 ( $P < 0.0001$ )	-0.69 ( $P < 0.0001$ )	-0.57 ( $P < 0.0001$ )	-0.49 ( $P = 0.0004$ )
Health insurance (1 000 000 yen)	0.2559	-0.40 ( $P = 0.0050$ )	-0.61 ( $P < 0.0001$ )	-0.46 ( $P = 0.0013$ )	-0.52 ( $P = 0.0002$ )	-0.49 ( $P = 0.0005$ )	-0.53 ( $P = 0.0001$ )	-0.61 ( $P < 0.0001$ )	-0.52 ( $P = 0.0002$ )
Elder ratio (%)	0.1377	-0.3928 ( $P = 0.0063$ )	-0.4804 ( $P = 0.0006$ )	-0.4441 ( $P = 0.0018$ )	-0.4768 ( $P = 0.0007$ )	-0.5824 ( $P < 0.0001$ )	-0.4456 ( $P = 0.0017$ )	-0.3794 ( $P = 0.0085$ )	-0.4291 ( $P = 0.0026$ )
Waste process (%)	0.3582	-0.39 ( $P = 0.0073$ )	-0.42 ( $P = 0.0030$ )	-0.43 ( $P = 0.0021$ )	-0.49 ( $P = 0.0005$ )	-0.54 ( $P = 0.0001$ )	-0.45 ( $P = 0.0017$ )	-0.36 ( $P = 0.0139$ )	-0.37 ( $P = 0.0100$ )

HFMD, Hand, foot and mouth disease; PCF, pharyngoconjunctival fever; GAS, group A streptococcus.

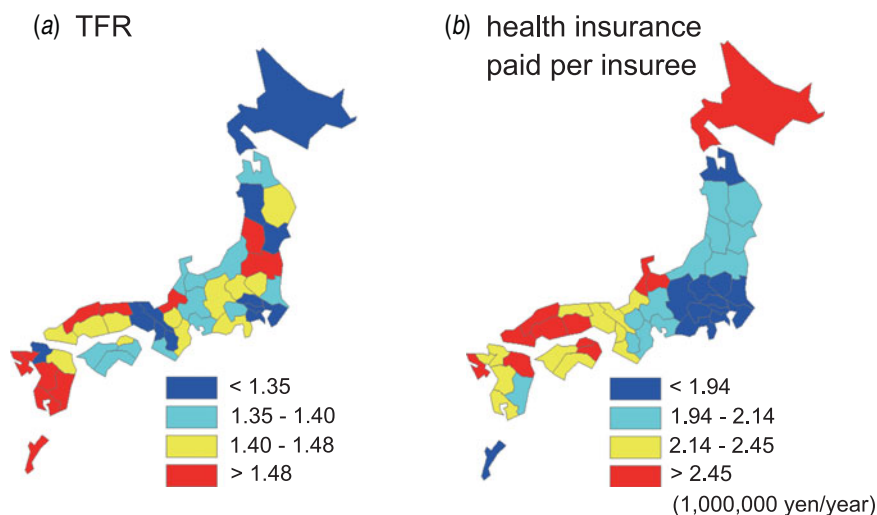
\* Uniqueness was estimated by factor analysis. A variable of smaller 'uniqueness' can be largely explained by a linear combination of the other variables.

Table 3. Conventional and spatial regression models to explain adjusted mean patient ages of Kawasaki disease and paediatric infectious diseases

Variable*	Kawasaki disease	Exanthema subitum	Herpangina	HFMD	Chickenpox	PCF	Erythema infectiosum	GAS
<b>Conventional regression</b>								
Total fertility rate	-0.29 ( <i>P</i> < 0.0001)	-0.13 ( <i>P</i> = 0.0207)	-1.9 ( <i>P</i> < 0.0001)	-1.7 ( <i>P</i> < 0.0001)	-0.92 ( <i>P</i> = 0.0053)	-1.2 ( <i>P</i> = 0.0019)	-0.69 ( <i>P</i> = 0.0159)	n.s.
Health insurance	n.s.	-1.3 ( <i>P</i> < 0.0001)	-3.1 ( <i>P</i> = 0.0228)	-4.2 ( <i>P</i> = 0.0020)	-2.2 ( <i>P</i> = 0.0151)	-8.0 ( <i>P</i> < 0.0001)	-6.2 ( <i>P</i> < 0.0001)	-11 ( <i>P</i> = 0.0001)
Varicella Vaccination	n.s.	n.s.	n.s.	n.s.	1.1 ( <i>P</i> = 0.0077)	n.s.	n.s.	n.s.
<i>R</i> <sup>2</sup>	0.34	0.45	0.56	0.58	0.62	0.46	0.44	0.29
<b>Spatial regression</b>								
Total fertility rate	-0.26 ( <i>P</i> < 0.0001)	-0.16 ( <i>P</i> = 0.0022)	-0.18 ( <i>P</i> < 0.0001)	-0.16 ( <i>P</i> < 0.0001)	-0.66 ( <i>P</i> = 0.0281)	-1.0 ( <i>P</i> = 0.0030)	-0.62 ( <i>P</i> = 0.0195)	n.s.
Health insurance	n.s.	-0.15 ( <i>P</i> < 0.0001)	n.s.	-2.7 ( <i>P</i> = 0.0253)	n.s.	-6.3 ( <i>P</i> = 0.0002)	-5.7 ( <i>P</i> < 0.0001)	-11 ( <i>P</i> = 0.0001)
Varicella Vaccination	n.s.	n.s.	n.s.	n.s.	1.4 ( <i>P</i> < 0.0001)	n.s.	n.s.	n.s.
<i>R</i> <sup>2</sup>	0.41	0.50	0.58	0.66	0.62	0.53	0.45	0.29

HFMD, Hand, foot and mouth disease; PCF, pharyngoconjunctival fever; GAS, group A streptococcus; n.s., not significant.

\* Although five variables (total fertility rate, pupils per class, health insurance, waste process, and chickenpox vaccination) were incorporated into the multivariate analysis, only total fertility rate, health insurance, and chickenpox vaccination remained as statistically significant contributor(s).



**Fig. 2.** Distribution of socioeconomic factors that were correlated with the mean patient age for each prefecture. The prefectures were categorized from red to blue in descending order for (a) the total fertility rate (TFR) and (b) the health insurance paid per insuree.

### Seasonality

We compared the seasonality across diseases. We presented our data starting in April to reflect the beginning of the Japanese academic school year (Fig. 3). Hand, foot, and mouth disease (HFMD) and herpangina exhibited a unimodal distribution with a sharp peak during the summer (July). GAS exhibited trimodal seasonality with three major peaks in June, December, and March. KD, chickenpox, pharyngoconjunctival fever (PCF), and erythema infectiosum exhibited a bimodal distribution with two major peaks, one in June and the other in December/January, as well as a subtle minor peak in March.

### Super-annual periodicity

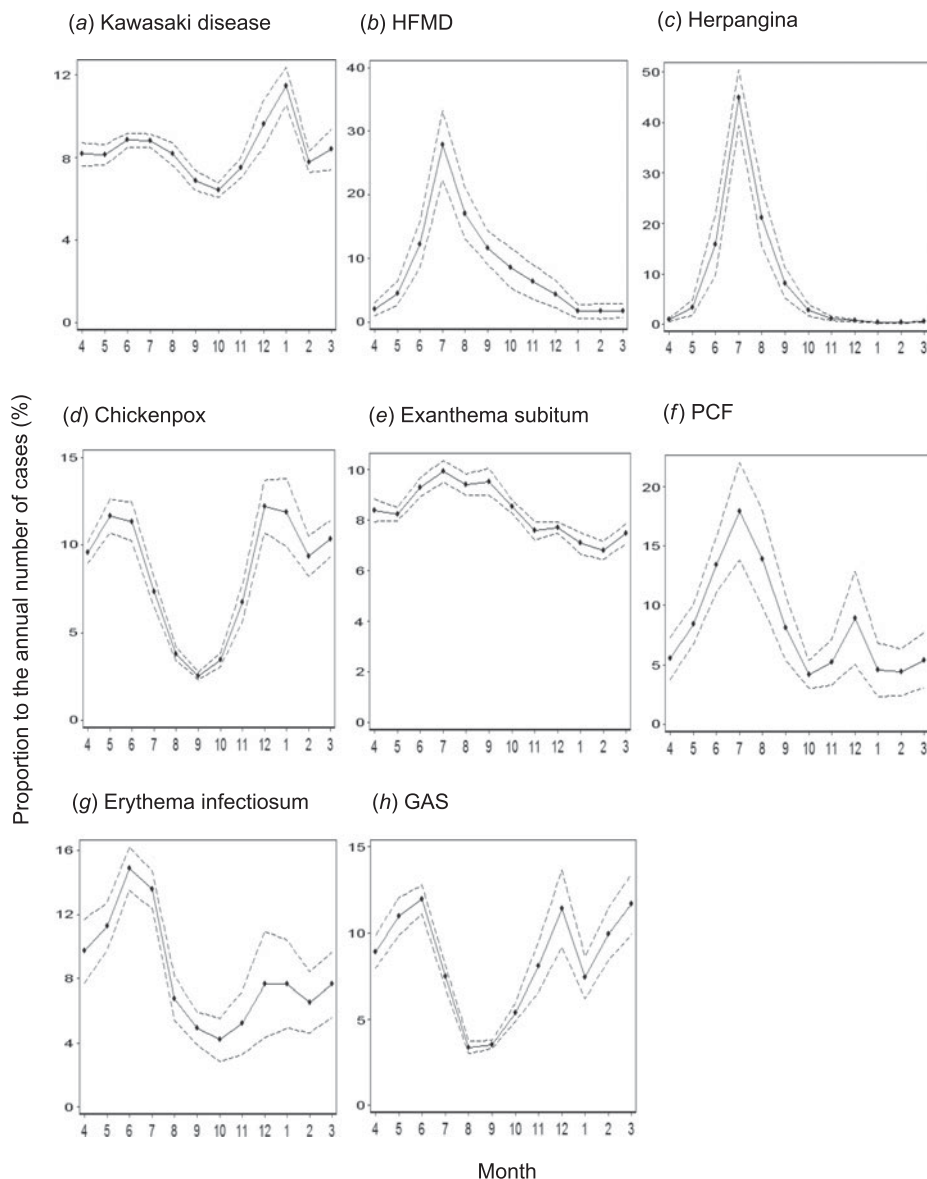
All the diseases examined in the present study, except KD and exanthema subitum, exhibited substantial fluctuations in the number of cases from year to year (Fig. 4). The wavelet analysis revealed that all the diseases except KD and exanthema subitum exhibited some coloured area, which implies that super-annual periodicity was detected in the corresponding period (Fig. 5). We obtained the incubation period ( $D$ ) and infectious period ( $D'$ ) of each disease from Richardson *et al.* [30] (Table 4) and used the adjusted mean patient age ( $A$ ) (Table 1) to test whether the super-annual periodicity could be explained by equation (7). Most of the super-annual periodicities estimated by the wavelet analysis were consistent with the predicted values (Table 4). We could not judge whether the prediction

for GAS was successful because  $D + D'$  was ambiguous for this disease. We also predicted  $D + D'$  for KD using equation (7). Because  $T$  was longer than 10 years after 1987 (Fig. 6) and  $A$  is 2.4 years, the  $D + D'$  for KD was estimated to be larger than 1.1 years. Since  $D$  for KD is approximately 10 days [31],  $D'$  for KD therefore is likely to be longer than 1 year. We will report the hypothesis elsewhere that the epidemics in 1979, 1982, and 1986 were driven by dynamics unique to these years (Y. Nagao, unpublished data).

### Probability of KD manifestation

Between 2006 and 2010, the average annual number of KD cases in Japan was 7699 for children aged between 0 and 2 years and 10 115 for children aged between 0 and 4 years. During the same period, the average number of live births was 1 082 997 per year and the death rate for children aged <5 years was 3667 per year. From these numbers, we estimated that approximately 0.7% and 0.9% of Japanese children had been diagnosed with KD by their third and fifth birthdays, respectively. By substituting  $A$  in equation (9) with the mean age of KD (i.e. 2.4 years), the proportion of individuals who experienced infection(s) with the aetiological agent for KD ( $1 - V$ ) was computed as 71% and 88% at  $t=3$  and  $t=5$  years, respectively. Therefore, only about 1% of infections with the KD agent(s) manifested with symptoms of KD.





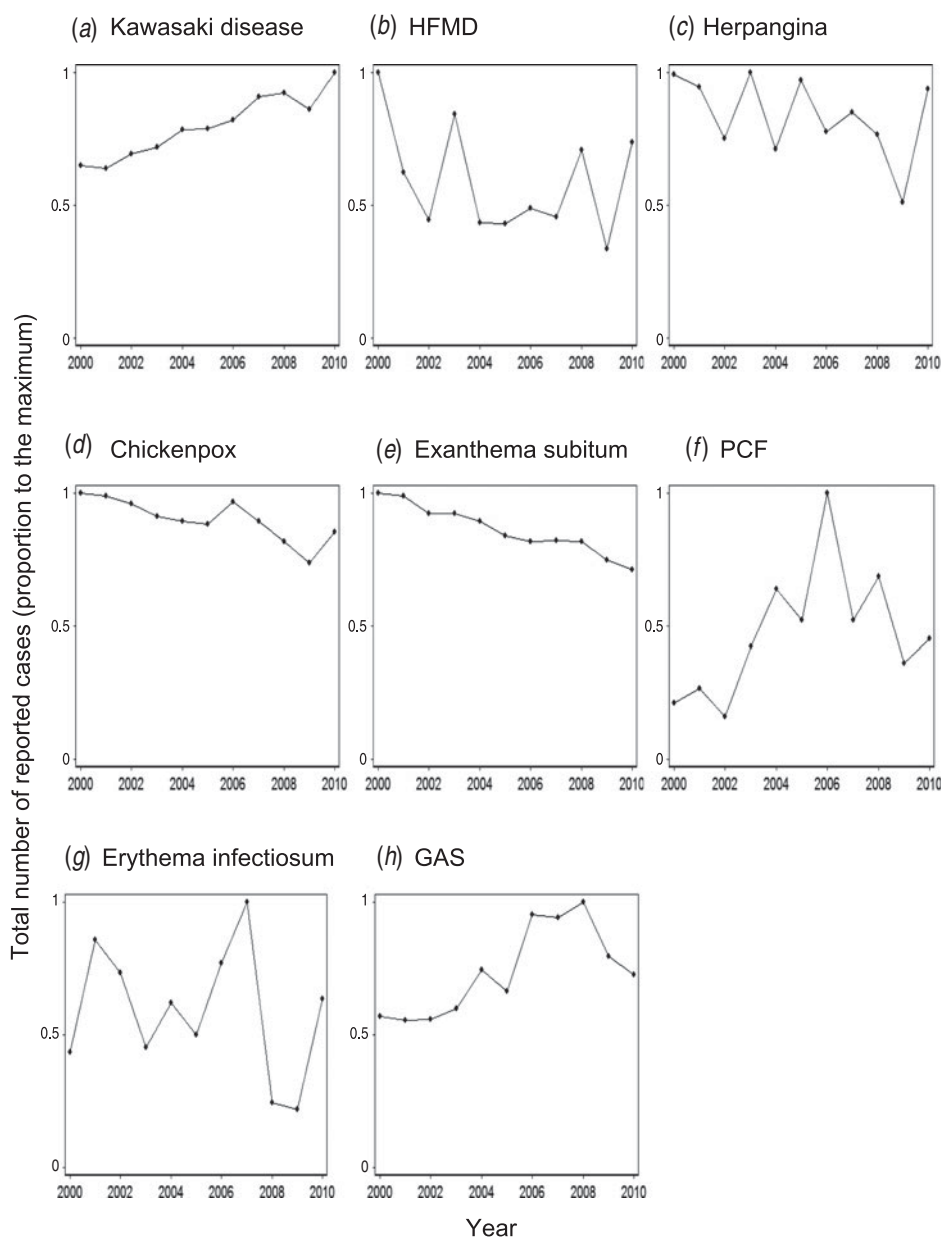
**Fig. 3.** Seasonality of Kawasaki disease and other paediatric infectious diseases. The proportion of the number of cases in a month to the annual number of cases was estimated for each year between 2000 and 2010 and averaged across these 11 years. The dashed lines indicate the standard deviation. The x-axis starts in April to reflect the beginning of the Japanese academic school year. HFMD, Hand, foot and mouth disease; PCF, pharyngoconjunctival fever; GAS, group A streptococcus.

## DISCUSSION

The geographical distribution of the FOI of KD and other paediatric infectious diseases was similar across Japan. We identified the availability/utilization of health services and the TFR as the main determinants of this shared geographical structure. Particularly, the TFR remained as a statistically significant risk factor to the FOI of KD and all the viral diseases, after accounting for the possible confounders and the spatial autocorrelation.

Indeed, there have been numerous reports that health services facilitate the transmission of paediatric infectious diseases [32, 33]. In addition, the causal relationship between the TFR and the FOI is intuitive: a higher TFR implies a larger number of siblings within a household, which increases the probability of sibling-to-sibling infection. Previous studies have also identified the importance of sibling-to-sibling infections for KD [31, 34].

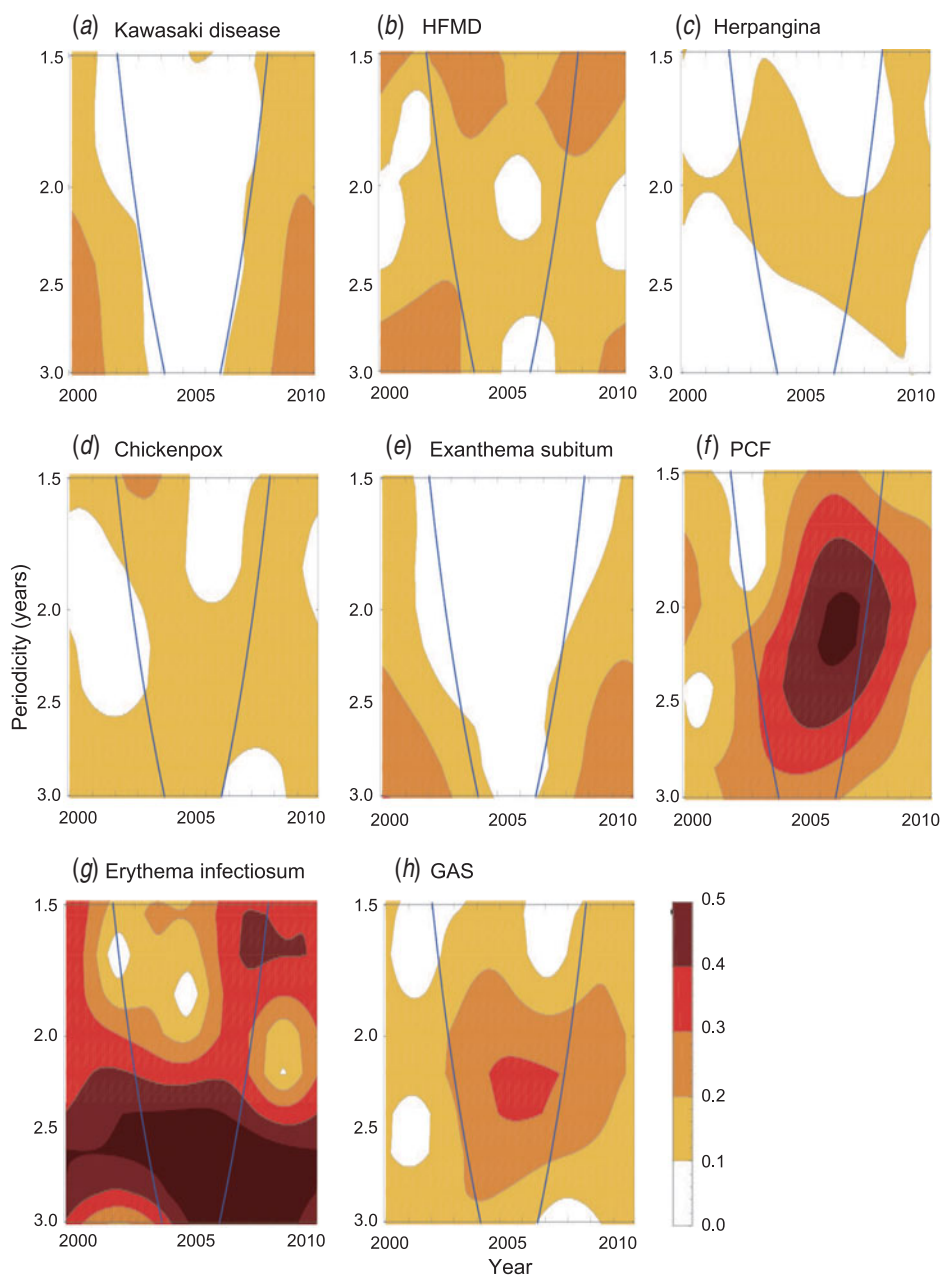
The finding that sibling-to-sibling infection and, to a lesser extent, hospital infection emerged as important



**Fig. 4.** Annual time series of Kawasaki disease (KD) and other paediatric infectious diseases in Japan between 2000 and 2010. The annual number of cases of KD and other paediatric infectious diseases was expressed as a proportion to the maximum number of reported cases. PCF, pharyngoconjunctival fever; GAS, group A streptococcus.

routes of transmission for KD and other paediatric infectious diseases in this study suggests that KD shares a mode of transmission with other paediatric infectious diseases, i.e. close contact with an infected individual. KD also exhibited a bimodal seasonality similar to other viral infections such as chickenpox and erythema infectiosum, each of which is caused by a single agent. Therefore, we cannot assert that KD is caused by multiple agents based solely on its multimodal seasonality. The seasonality of a paediatric infectious disease is

largely determined by climate and the school term [35, 36]. Indeed, the troughs in the seasonality of GAS occurred in April, September, and January, immediately following school/kindergarten holidays in the spring (late March to early April), the summer (late July to late August), and the winter (late December to early January), respectively. This finding suggests that these troughs were due to the separation of children during the holidays. This finding also suggests that the bimodal seasonalities of KD, chickenpox,



**Fig. 5.** Super-annual periodicities of KD and other paediatric infectious diseases as detected by the wavelet analysis. The blue lines indicate the 95% confidence limit of the wavelet analysis. The area outside the blue lines is unreliable due to the edge effect. PCF, pharyngoconjunctival fever; GAS, group A streptococcus.

PCF, and erythema infectiosum were affected by the school term. Our finding that KD shares a similar seasonality pattern to other paediatric diseases suggests that the transmission routes for these diseases are common. Taken together, these results suggest that KD is not likely to be transmitted through a novel route such as a transcontinental wind [37, 38].

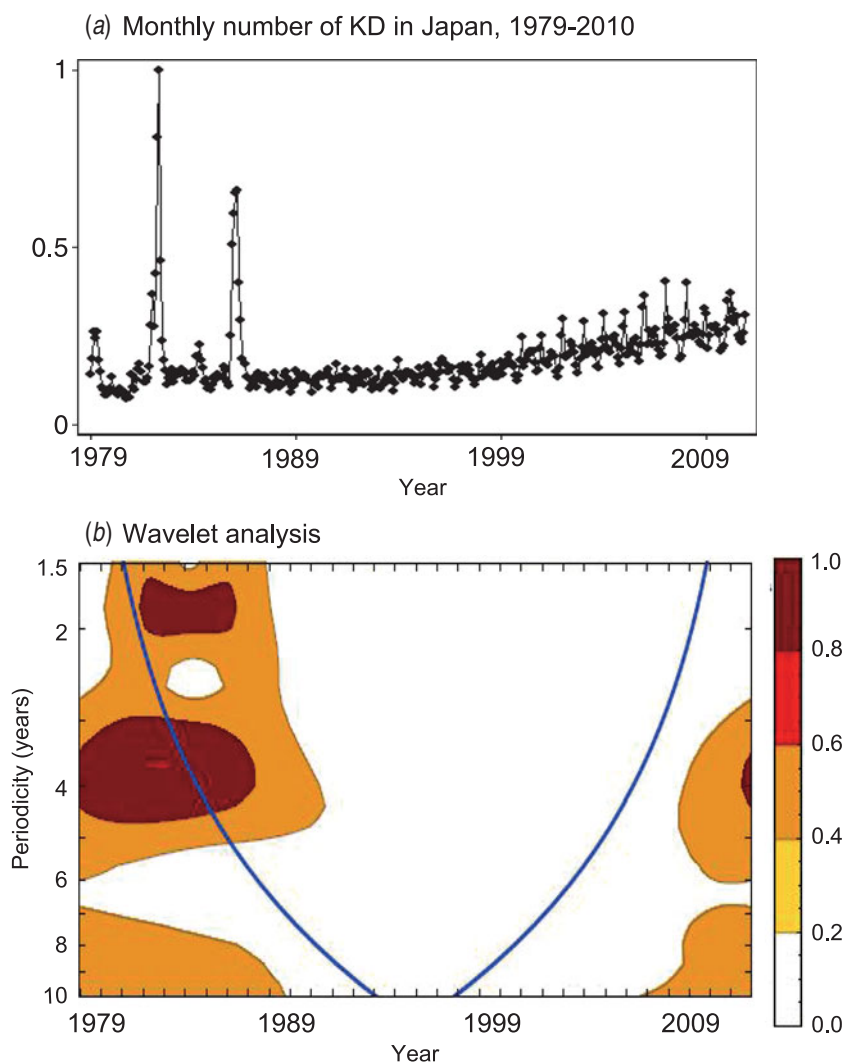
One notable outlier in our regression analysis was Okinawa prefecture, a tropical prefecture composed of

160 islands located 650 km from Japan's mainland. Host genetic heterogeneity is an unlikely explanation for this observation, because the genetic structure of the Okinawan population is similar to other populations within Japan [39]. The tropical climate of Okinawa may contribute to this observation [40]. Alternatively, due to the remoteness and multi-island structure of Okinawa, the transient extinction of a pathogen may be more frequent than on the mainland of Japan [41].

Table 4. *The lengths of predicted and actual super-annual periodicities (T)*

	Incubation period (D)	I Infectious period (D')	$D + D'$	T predicted by equation (2)	T measured in actual time series by wavelet analysis
Exanthema subitum	10–15 d	Lifelong	>10 yr	>20 yr	>3.0 yr
Herpangina	3–5 d	<7 d	11 d	2.0 yr	2.2 yr
HFMD	3–5 d	<7 d	11 d	2.1 yr	1.7 yr
Chickenpox	11–20 d	–1 to +2 d	16 d	2.5 yr	>2.0 yr
PCF	3–29 d	<2 wk	16 d	2.8 yr	2.2 yr
Erythema infectiosum	13–18 d	–6 d to –3 d	9 d	2.4 yr	2.6 yr
GAS	12 h–5 d	(1) untreated: 1–12 mo. (2) treated: <4 d	(1) 3 mo. (2) 8 d	(1) 8.3 yr (2) 2.5 yr	2.3 yr

HFMD, Hand, foot and mouth disease; PCF, pharyngoconjunctival fever; GAS, group A streptococcus; d, days; wk, weeks; mo., months; yr, years.



**Fig. 6.** Monthly number of Kawasaki disease (KD) cases between 1979 and 2010. The monthly number of KD cases recorded between 1979 and 2010 in Japan are expressed as the proportion to the maximum number of cases (a). The results of the wavelet analysis applied to this time series are plotted in (b).

One major limitation of the present study was that the data about the paediatric infectious diseases was derived from sentinel paediatric clinics. The allocation of sentinel clinics is not precisely proportional to the population size. However, no factor was suggested as a possible source of bias in estimating the mean patient age from the sentinel data. Therefore, we believe that our analyses were not affected significantly by the use of sentinel data.

Another possible refutation against the present analysis is that our statistical analysis may have been flawed by ecological confounding. However, we believe that the identified contributions of fertility rate and hospital utilization to the FOI are not statistical artefacts. Four of the eight variables, which were pre-selected from a large number of variables during the initial screening process, were related to the utilization of health services (Table 2). This result strongly suggests that health services play an important role in determining the FOI. In addition, the varicella vaccine uptake rate remained as a statistically significant predictor of risk reduction on the FOI of chickenpox only (Table 3). This result validates our analytical procedure, because vaccination is assumed to decrease the FOI [21, 42]. Moreover, the fertility rate (or birth rate) is one of the most important factors that determine the epidemiology of paediatric infectious diseases [28].

Our mathematical assumptions may seem to contain excessive approximations. For example, equation (5) presumes a fixed longevity (or Type I survival). This, however, is a well-founded approximation. First, Type I survival is more suitable 'for humans, especially in developed countries' than Type II survival, which assumes age-independent mortality [21]. Second, regardless of whether Type I or II survival is assumed, the mean patient age can effectively approximate the inverse of the FOI, if the mortality is very small as in Japan. Equation (8), which has been used in epidemiological studies [43], presumes continuous susceptibility to a given disease from birth and complete endemicity of the disease agent. Therefore, our prediction based upon equation (8) should be tested in future serological surveys once the aetiological agent for KD is identified.

Given that any infectious disease has finite values for the incubation and infectious periods (i.e.  $D + D'$ ), equation (7) predicts that any infectious disease will have a finite super-annual periodicity ( $T$ ). However, among the diseases examined in this study, only KD and exanthema subitum exhibited the least evidence

for super-annual periodicities. This finding suggests that the super-annual periodicities of these diseases are very long, and hence the sum of the incubation period and infectious period of these diseases is large [equation (7)]. Consistent with this finding, the viral agent for exanthema subitum persists across the life-span [44]. Therefore, it would not be surprising if the aetiological agent for KD also persists for a long period. Our finding is consistent with a previous report that the agent (or at least one of the agents) for KD is capable of persistent infection [19].

The low mean patient age, and therefore the strong FOI of KD, led to an estimation that only about 1% of infections with the KD agent(s) manifest with symptoms of KD. Our finding that KD may persist over a long period and cause frequent asymptomatic infections predicts that the aetiological agent(s) for KD exists in KD patients as well as in many healthy individuals. Therefore, an epidemiological study to determine whether a particular microbe is a causative agent for KD will require a large sample size to detect subtle difference in the prevalence of the microbe between cases and controls. In addition, the possibility that sibling-to-sibling infection is an important route of transmission for KD implies that the number of siblings and the rank of the patient among them should be considered as covariates in epidemiological analyses.

## SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit <http://dx.doi.org/10.1017/S0950268815001223>.

## ACKNOWLEDGEMENTS

We are grateful to Yoshikazu Nakamura and Mayumi Yashiro for providing the KD data. We are also grateful to Maurizio Pisati and Junji Shiraishi, who assisted in the spatial analysis. We thank Tetsuya Sano for helpful comments on the manuscript.

This research was supported in part by the Aihara Project, the FIRST programme from the Japan Society for the Promotion of Science, which was initiated by the Council for Science, Technology, and Innovation.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **Kawasaki T.** Acute febrile mucocutaneous syndrome with lymphoid involvement with specific desquamation of the fingers and toes in children [in Japanese]. *Japanese Journal of Allergology* 1967; **16**: 178–222.
2. **Nakamura Y, et al.** Increasing incidence of Kawasaki disease in Japan: nationwide survey. *Pediatrics International* 2008; **50**: 287–290.
3. **Burns JC, et al.** Seasonality and temporal clustering of Kawasaki syndrome. *Epidemiology* 2005; **16**: 220–225.
4. **Nakamura Y, Yanagawa I, Kawasaki T.** Temporal and geographical clustering of Kawasaki disease in Japan. *Progress in Clinical and Biological Research* 1987; **250**: 19–32.
5. **Shingadia D, Bose A, Booy R.** Could a herpesvirus be the cause of Kawasaki disease? *Lancet Infectious Diseases* 2002; **2**: 310–313.
6. **Tsuchida S, et al.** Epidemiology of infant Kawasaki disease with a report of the youngest neonatal case ever reported in Japan. *Acta Paediatrica* 1996; **85**: 995–997.
7. **Hirata S, Nakamura Y, Yanagawa H.** Incidence rate of recurrent Kawasaki disease and related risk factors: from the results of nationwide surveys of Kawasaki disease in Japan. *Acta Paediatrica* 2001; **90**: 40–44.
8. **Burgner D, et al.** A genome-wide association study identifies novel and functionally related susceptibility Loci for Kawasaki disease. *PLoS Genetics* 2009; **5**: e1000319.
9. **Khor CC, et al.** Genome-wide linkage and association mapping identify susceptibility alleles in ABCC4 for Kawasaki disease. *Journal of Medical Genetics* 2011; **48**: 467–472.
10. **Onouchi Y, et al.** A genome-wide association study identifies three new risk loci for Kawasaki disease. *Nature Genetics* 2012; **44**: 517–521.
11. **Kato H, et al.** Coronary aneurysms in infants and young children with acute febrile mucocutaneous lymph node syndrome. *Journal of Pediatrics* 1975; **86**: 892–898.
12. **Kato H, et al.** Long-term consequences of Kawasaki disease. A 10- to 21-year follow-up study of 594 patients. *Circulation* 1996; **94**: 1379–1385.
13. **Furusko K, et al.** High-dose intravenous gammaglobulin for Kawasaki disease. *Lancet* 1984; **324**: 1055–1058.
14. **Ogata S, et al.** Coronary artery outcomes among children with Kawasaki disease in the United States and Japan. *International Journal of Cardiology* 2013; **168**: 3825–3828.
15. **Amaku M, et al.** Relationship among epidemiological parameters of six childhood infections in a non-immunized Brazilian community. *Memorias do Instituto Oswaldo Cruz* 2009; **104**: 897–900.
16. **Nationwide Surveillance of Kawasaki Disease [Kawasaki byo zenkoku chosa].** (<http://www.jichi.ac.jp/dph/kawasaki.html>). Accessed 13 December 2014.
17. **Nagao Y.** Decreasing fertility rate correlates with the chronological increase and geographical variation in incidence of Kawasaki disease in Japan. *PLoS ONE* 2013; **8**: e67934.
18. **Annual report of infectious diseases [Kansensho hassei doko chosa jigyo nenpo].** (<http://idsc.nih.go.jp/idwr/CDROM/Main.html>). Accessed 3 September 2012.
19. **Pitzer VE, et al.** Modelling seasonal variations in the age and incidence of Kawasaki disease to explore possible infectious aetiologies. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 2012; **279**: 2736–2743.
20. **Nagao Y, et al.** Geographical structure of dengue transmission and its determinants in Thailand. *Epidemiology and Infection* 2008; **136**: 843–851.
21. **Anderson RM, May RM.** *Infectious Diseases of Humans: Dynamics and Control*. New York: Oxford University Press, 1991.
22. **Yu AL, et al.** Varicella transmission in two samples of children with different social behaviour in the State of Sao Paulo, Brazil. *Epidemiology and Infection* 2001; **127**: 493–500.
23. **Nagao Y, Koelle K.** Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever. *Proceedings of the National Academy of Sciences USA* 2008; **105**: 2238–2243.
24. **Global summary of day/month observations, 1979–continuing (ds512.0)** (<http://dss.ucar.edu/datasets/ds512.0/>). Accessed 20 September 2012.
25. **Statistical indicators for social life [Shakai seikatsu tokei shihyou].** (<http://www.e-stat.go.jp>). Accessed 8 March 2013.
26. **Anselin L, Hudak S.** Spatial econometrics in practice. A review of software options. *Regional Science and Urban Economics* 1992; **22**: 509–536.
27. **Hay SI, et al.** Etiology of interepidemic periods of mosquito-borne disease. *Proceedings of the National Academy of Sciences USA* 2000; **97**: 9335–9339.
28. **Grenfell BT, Bjornstad ON, Kappey J.** Travelling waves and spatial hierarchies in measles epidemics. *Nature* 2001; **414**: 716–723.
29. **Torrence C, Compo GP.** A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* 1998; **79**: 61–78.
30. **Richardson M, et al.** Evidence base of incubation periods, periods of infectiousness and exclusion policies for the control of communicable diseases in schools and preschools. *Pediatric Infectious Disease Journal* 2001; **20**: 380–391.
31. **Fujita Y, et al.** Kawasaki disease in families. *Pediatrics* 1989; **84**: 666–669.
32. **Johnston JM, Burke JP.** Nosocomial outbreak of hand-foot-and-mouth disease among operating suite personnel. *Infection Control* 1986; **7**: 172–176.
33. **Buettcher M, Heining U.** Prospective surveillance of nosocomial viral infections during and after hospitalization at a university children's hospital. *Pediatric Infectious Disease Journal* 2010; **29**: 950–956.
34. **Tsai HC, et al.** Transmission of acute infectious illness among cases of Kawasaki disease and their household members. *Journal of the Formosan Medical Association* 2014.
35. **Brisson M, et al.** Epidemiology of varicella zoster virus infection in Canada and the United Kingdom. *Epidemiology and Infection* 2001; **127**: 305–314.
36. **Grassly NC, Fraser C.** Seasonal infectious disease epidemiology. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 2006; **273**: 2541–2550.

37. **Rodo X, et al.** Association of Kawasaki disease with tropospheric wind patterns. *Scientific Reports* 2011; **1**: 152.
38. **Rodo X, et al.** Tropospheric winds from northeastern China carry the etiologic agent of Kawasaki disease from its source to Japan. *Proceedings of the National Academy of Sciences USA* 2014; **111**: 7952–7957.
39. **Yamaguchi-Kabata Y, et al.** Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *American Journal of Human Genetics* 2008; **83**: 445–456.
40. **Burns JC, et al.** Seasonality of Kawasaki disease: a global perspective. *PLoS ONE* 2013; **8**: e74529.
41. **Nodelijk G, et al.** Introduction, persistence and fade-out of porcine reproductive and respiratory syndrome virus in a Dutch breeding herd: a mathematical analysis. *Epidemiology and Infection* 2000; **124**: 173–182.
42. **Laurenson MK, et al.** Prevalence, spatial distribution and the effect of control measures on louping-ill virus in the Forest of Bowland, Lancashire. *Epidemiology and Infection* 2007; **135**: 963–973.
43. **Thai KT, et al.** Incidence of primary dengue virus infections in Southern Vietnamese children and reactivity against other flaviviruses. *Tropical Medicine and International Health* 2007; **12**: 1553–1557.
44. **Suga S, et al.** Prospective study of persistence and excretion of human herpesvirus-6 in patients with exanthem subitum and their parents. *Pediatrics* 1998; **102**: 900–904.