Epidemiol. Infect. (2015), **143**, 522–528. © Cambridge University Press 2014
doi:10.1017/S0950268814000697

The spatial diffusion of norovirus epidemics over three seasons in Tokyo

S. INAIDA[1]*, Y. SHOBUGAWA[1], S. MATSUNO[2], R. SAITO[1] AND H. SUZUKI[3]

[1] *Division of International Health (Public Health), Graduate School of Medical and Dental Sciences, Niigata University, Niigata, Japan*
[2] *Infectious Disease Surveillance Centre, National Institute of Infectious Diseases, Tokyo, Japan*
[3] *School of Nursing, Niigata Seiryo University, Niigata, Japan*

Received 14 September 2013; Final revision 2 March 2014; Accepted 9 March 2014; first published online 29 April 2014

SUMMARY

We studied the spatial trend of norovirus (NoV) epidemics using sentinel gastroenteritis surveillance data for patients aged <15 years ($n=140$) in the Tokyo area for the 2006–2007 to 2008–2009 seasons utilizing the kriging method of geographical information system (GIS). This is the first study of the spreading pattern of NoV epidemics using sentinel surveillance data. Correlations of sentinel cases between the seasons and with demographic data were examined to identify the trend and related factors. A similar pattern of diffusion was observed over the seasons, and its mean correlation between seasons was significantly high. A higher number of cases were found in the peripheral area, which surrounds the most populated central area, and showed a correlation with the ratio of the children population ($r=0{\cdot}321$, $P<0{\cdot}01$) and the ratio of residents in larger families ($r=0{\cdot}263$, $P<0{\cdot}01$). While NoV susceptibility remained, the results suggest a transmission route in the local community as a possible epidemic factor. Prevention with focus on the peripheral area is desirable.

Key words: Geographical information system, household, kriging method, national sentinel surveillance, norovirus, population.

INTRODUCTION

Norovirus (NoV) is the most frequent cause of viral gastroenteritis with considerable morbidity and mortality annually [1–6]. Epidemics occur from the autumn to winter seasons in Japan as well as other temperate climate regions [7]. NoV patients present at varying ages, although severe cases are more commonly found in young children or patients with chronic diseases, and may result in death [3, 5–9]. NoV is highly transmissible through food, water, and person-to-person contact [10, 11]. The virus is spread by the faecal–oral route through inhalation of virus particles from vomit and faeces or contact with contaminated surfaces [12–16]. Frequent molecular evolutions often escape immunity such that repeated infection can occur [8, 17–20]. The role of environmental contamination is suggested in continued or prolonged epidemics [8, 9, 11, 17, 21–23]. Seasonal epidemics raise the question about how epidemic spread occurs in the community. The discovery of a pattern of epidemics is expected to lead to models of prevention or timely preparedness. Currently, surveillance data are mostly designed for conventional statistical use; however, a spatial study requires

* Author for correspondence: Dr S. Inaida, Division of International Health, Graduate School of Medical and Dental Sciences, Niigata University, 1-757, Asahimachi-dori, Chuo-ku, Niigata City, Niigata, 951-8510, Japan.
(Email: inaida@med.niigata-u.ac.jp)

spatiotemporal data collection. Thus, few studies have been performed regarding the spread of NoV epidemics. Here we attempt to apply sentinel surveillance data into a spatial method and consider the pattern of epidemic spread.

In Japan, NoV epidemics are monitored through sentinel gastroenteritis surveillance that is routinely operated year round. The largest NoV epidemic in the history of national surveillance since its launch in 1981 occurred in the 2006–2007 season. This epidemic was the result of the emergence of a GII.4 subtype termed 2006b [24]. We investigated the spatial diffusion of NoV epidemics in the Tokyo area, which is the most populated area in Japan, using sentinel gastroenteritis data applied to the geographical information system (GIS). Because we assumed characteristic patterns exist in comparisons between the largest epidemic season and the following seasons, we therefore focused on the 2006–2007 to 2008–2009 seasons. GIS is increasingly being used for disease surveillance and is useful when considering the spatial spread of an epidemic and its related factors by using other background data such as environmental or community data. We examined the relationships between sentinel reported cases and demographic data to identify factors that may influence virus transmission.

MATERIALS AND METHODS

Sentinel gastroenteritis data from the 2006–2007 to 2008–2009 seasons were used to investigate NoV epidemic trends in these three seasons, including consideration of whether the epidemics occur in relation to a populated area and also to observe if there occurred any epidemic features over the seasons after a large outbreak. Sentinel gastroenteritis data were extracted for the NoV season (September to December) [25], and the weekly flow of data and the age distribution of the total number of sentinel cases were compared among the seasons. The map of epidemic spread was depicted by means of the kriging method in GIS applied to the total number of sentinel cases. Then the pattern of epidemic spread and demographic data were considered. Correlations of the total number of sentinel cases between seasons were also examined. In order to examine the relationship between the epidemics and the demographic characteristics, a spatial overlay of each data was investigated. For this analysis, the demographic data were extracted using a 1 km radius for each sentinel site. This circle was chosen because most sentinel sites were small to medium paediatric clinics, therefore, the cases reported by a sentinel clinic are assumed to come from a circular catchment area of radius 1 km (which is about a 15 min walking distance) centred at the location of the sentinel site. Pearson's correlation and the coefficient between the total number of sentinel cases and the extracted demographic data were evaluated for each demographic index, respectively.

Sentinel surveillance data

Sentinel surveillance data were obtained from the National Epidemiological Surveillance of Infectious Diseases (NESID) of the National Institute of Infectious Diseases (NIID), which comprises weekly reporting of the number of cases from about 3000 sentinel clinics. The sentinel clinics for gastroenteritis are paediatric clinics, meaning the majority of data consist of cases of patients aged <15 years. We accordingly focused our study on only sentinel surveillance data by extracting data for patients aged 0–14 years. The data were first collected by a physician who diagnosed gastroenteritis based on symptoms (diarrhoea, vomiting, acute abdominal pain) and then reported to NESID through regional health centres. The flow of surveillance and the reporting system are described elsewhere [26]. The number of sentinel clinics is defined by the population size according to each regional health centre's administrative sector, and the average number of reported cases per sentinel site is used as an index of the epidemic level of the season. In Tokyo, the number of sentinel clinics for gastroenteritis is 140. NoV epidemics were also monitored by laboratory confirmation [26, 27]. For laboratory testing, specimens were collected from about 10% of the sentinel clinics and on-site for any outbreaks. Viral genotyping was performed by means of reverse transcriptase–polymerase chain reaction methods by prefectural institutes, and the results reported to NESID. These data include both sporadic cases and outbreaks.

The data period was epidemiological weeks 36–52 (the beginning of September to the end of December). This period was defined as the season of the NoV epidemic based on results of historical viral detection and the trend of sentinel reported cases [28].

Trend of incidence

The weekly distribution of the number of sentinel cases (per sentinel site) as incidences was

https://doi.org/10.1017/S0950268814000697 Published online by Cambridge University Press

examined to identify the epidemic trend for the three seasons.

Total incidences of 8 weeks until the peak week (henceforth 'total number of sentinel cases') for each season were calculated from sentinel surveillance data. These 8 weeks approximate the period when the sentinel case starts to relatively rapidly increase (over 5·0 cases per sentinel per week) in the three seasons and include the 4th week of October to the 2nd week of December in the 2006–2007 season, the 5th week of October to the 4th week of December in the 2007–2008 season, and the 4th week of October to the 2nd week of December in the 2008–2009 season. The total number of sentinel cases was used as an index of the epidemic level for each season while taking account of the growth of the epidemic until the peak week. The total number of sentinel cases for each sentinel site was also calculated and correlations between the three seasons were considered for 140 sentinel sites. The cases' age distribution was also calculated for these seasons from sentinel data.

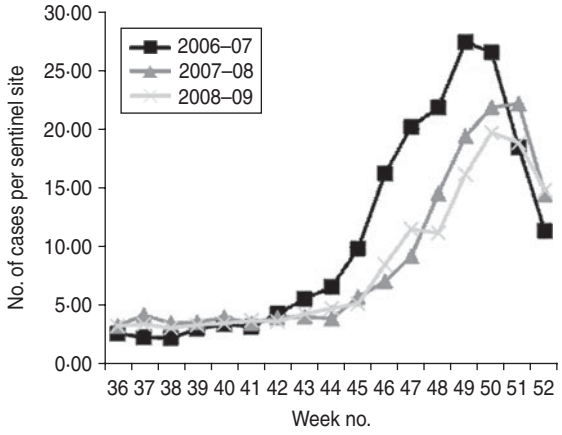

Fig. 1. Weekly distribution of norovirus sentinel cases for three seasons (2006–2007, 2007–2008, 2008–2009) in Tokyo.

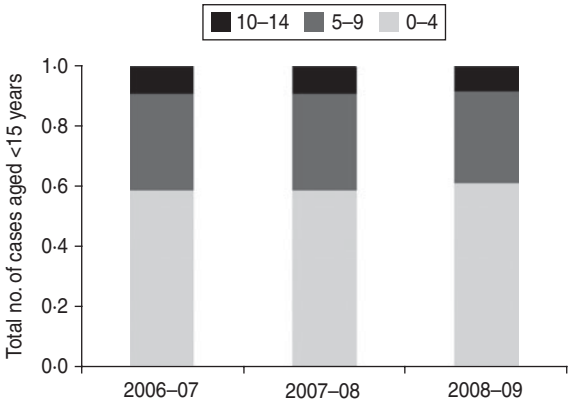

Fig. 2. Age group distribution of norovirus sentinel cases for three seasons (2006–2007, 2007–2008, 2008–2009) in Tokyo.

Spatial analysis with demographic data

Observation of epidemic spreads and demographic data

The spatial spread of epidemic diffusion was considered using the kriging method (ordinary kriging) in GIS (ArcGIS v. 9.3, ESRI, Japan) which was applied to the total number of sentinel cases at 140 sentinel sites for each of the three seasons. The kriging method is an interpolation method that uses geostatistical estimation over point-based data and models the surface of each record [29]. In the present study, the records are sentinel surveillance data. This method enabled the investigation of areas with a higher number of cases. Based on the assumption that epidemics are affected by population characteristics, we mapped the demographic data for the ratio of the population aged <15 years to the total population and for the ratio of residents in a household having $\geqslant 3$ family members to the total population as gridded 1-km maps. The demographic data came from the national census (2005).

Correlations between sentinel cases and demographic indices

Pearson correlation coefficients between the total number of sentinel cases of the 2006–2007 season and the demographic data extracted by a 1 km radius for each sentinel site were examined, respectively. The demographic data were obtained from the 2005 national census and consisted of (*a*) population indices, which included the total population, children by age group (0–4, 5–9, 10–14, $\geqslant 15$ years) and the corresponding ratios of these age groups to the total population, and (*b*) household indices, which included the number of households, the ratio of each house style (home or apartment) to the total number of house styles, and the number of residents in a household having 1, 2, or $\geqslant 3$ family members along with the ratio of these groups to the total population. House style was considered because it might relate to living characteristics causing contact development inside or outside home. Statistical analysis was conducted using SPSS (IBM Corporation, USA).

https://doi.org/10.1017/S0950268814000697 Published online by Cambridge University Press

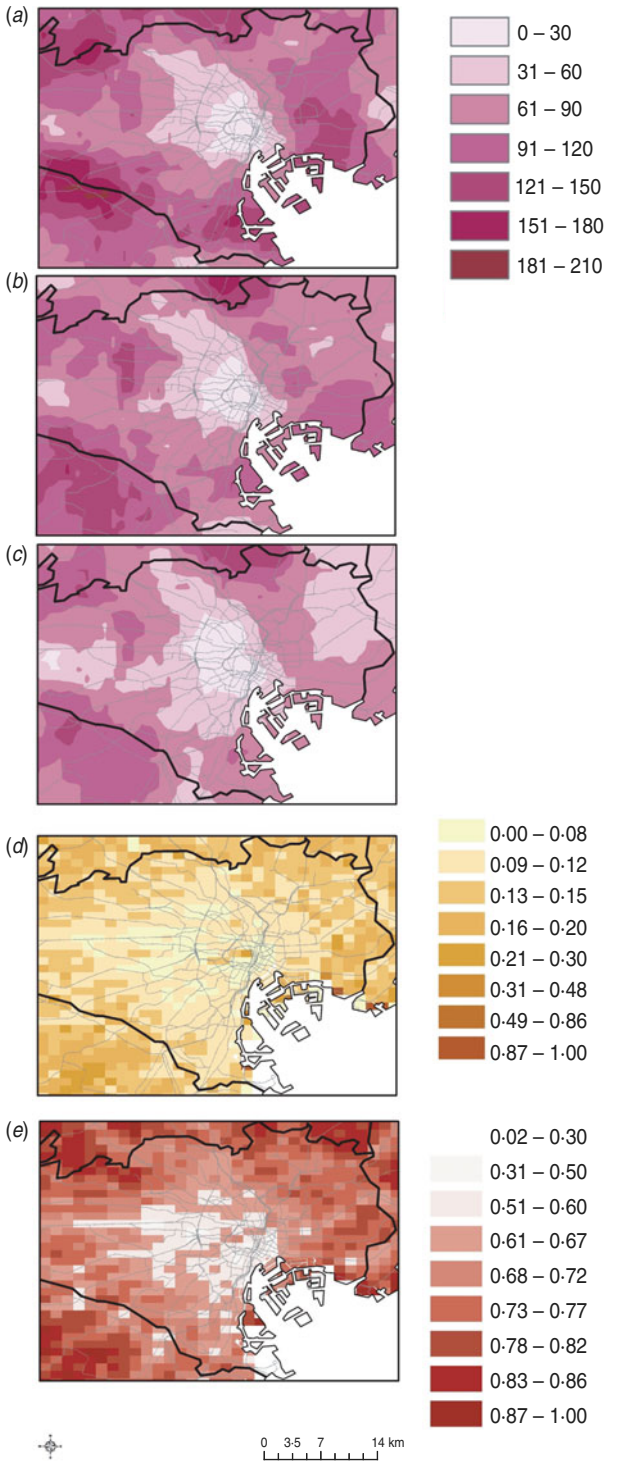

Fig. 3. Spatial spread of norovirus (NoV) sentinel cases and demographic data. Spread of the total number of NoV sentinel cases (from 8 weeks prior to the peak week) depicted by the kriging method in GIS for the (*a*) 2006–2007, (*b*) 2007–2008, and (*c*) 2008–2009 seasons. These maps show the geographical concentration of the epidemic in each season. Demographic data are shown in gridded 1-km maps as (*d*) the ratio of children (aged <15 years) compared to the total population and (*e*) the ratio of residents in households having ⩾3 family members compared to the total population. Bold black lines signify the prefecture boundaries; grey lines signify railways. The unit values in panels (*a–c*) are the total number of sentinel cases. The unit values in panels (*d*) and (*e*) are the ratios between the number of the corresponding index population and the total population. All classification colours are in ascending order. Original map data (ESRI, Japan) are based on national census data (2005).

RESULTS

Trends in incidence and spread of epidemics

The weekly incidences showed the highest peak at 27·44 cases in the 2006–2007 season, 22·17 cases in the 2007–2008 season, and 19·72 cases in the 2008–2009 season (per sentinel site). The total number of sentinel cases recorded was 112·09 in the 2006–2007 season, 104·00 in the 2007–2008 season, and 81·29 in the 2008–2009 season (Fig. 1). The incidences at the peak week and the total number of sentinel cases slightly decreased over the seasons, which may suggest the development of herd immunity after a period of high prevalence. Within the incidences, each age group showed similar ratios over the three seasons. The mean value of the three seasons was 0·59 for the 0–4 years age group, which was the highest recorded, 0·32 for the 5–9 years age group, and 0·09 for the 10–14 years age group (Fig. 2).

The epidemic diffusion depicted by the kriging method demonstrated areas of different concentrations of total number of sentinel cases (Fig. 3*a–c*). A larger number of data were observed around the peripheral area, which surrounded the central core area, and the central core area had a similar lower number of data for the three seasons. Correlations of the total number of sentinel cases between the 2006–2007 season and other seasons (at 140 sentinel sites) were significantly positive, reaching 0·89 for the 2007–2008 season ($P<0·01$), and 0·87 to the 2008–2009 season ($P<0·01$). These results showed similar patterning of the epidemic spread over the seasons.

The demographic data showed that the peripheral area had a higher ratio of residents in households with ⩾3 family members and a higher ratio for those aged <15 years compared to the total population (Fig. 3*d, e*). These data showed the peripheral area as a suburban area with more residents in larger households than residents in households having one or two family members; and similarly, more residents in families with children, compared to the central area,

https://doi.org/10.1017/S0950268814000697 Published online by Cambridge University Press

Table 1. *Correlation between total number of sentinel cases (2006–2007 season) and demographic data (N = 140)*

Population indices		Household indices	
Total population	−0·037	No. of households	−0·076
Age group population, yr		Ratio of house style	
0–4	0·145	Home	0·163
5–9	0·147	Apartment	−0·163
10–14	0·120	No. of family members	
⩾15	0·140	1–2	−0·106
Ratio of age group population to total population		⩾3	0·048
0–4	0·305**	Ratio of family members to total population	
5–9	0·318**	1–2	−0·263**
10–14	0·284**	⩾3	0·263**
⩾15	0·321**		

Demographic data are extracted within a 1 km radius from each sentinel.
Original map data (ESRI, Japan) are based on national census data (2005).
**$P < 0·01$.

which consisted of a higher density of population. In the central area, a higher density of children was also observed (Supplementary Fig. S1 only shows children data but population were similarly concentrated in the central area). Thus lower density but higher ratio of larger families or children seemed to be one of the community characteristics in the peripheral area.

Correlations between sentinel data and demographic data

The total number of sentinel cases significantly correlated to the following population ratios: the ratios of children of all age groups (0–4, 5–9, 10–14, ⩾15 years) to the total population (0·284–0·321, $P < 0·01$), and the ratio of residents in households having ⩾3 family members to the total population (0·263, $P < 0·01$) (Table 1). However, no significant correlations were seen for the absolute numbers of the total population, children of each age group (0–4, 5–9, 10–14, ⩾15 years), or of each household size. Neither did the ratio of each house style (home or apartment) to the number of all house styles show any correlation.

These results suggest a relationship between an increase in the epidemic with an area with a relatively larger component of larger families or children, rather than with an area with a larger number of population of all age groups or in larger families, but with a larger component of smaller families.

DISCUSSION

The total number of sentinel cases showed a larger number around the peripheral area of Tokyo, which includes the area surrounding the central area. Pearson correlation coefficients were significantly positive between the total number of sentinel cases and indices for the ratio of children of all age groups (0–4, 5–9, 10–14, ⩾15 years) and the ratio of residents in households with ⩾3 family members. However, there was no significant correlation with other indices like the total population or the absolute number of children of each age group or household.

Although it is thought that NoV would develop larger incidences in a densely populated community by increasing person-to-person contact, our results show otherwise, as it was the peripheral community areas, which have a lower density of population or children, but a higher ratio of children and residents in larger families, that correlated with NoV epidemics. A possible explanation for why the peripheral area had more patients is, hypothetically, more frequent contacts and communications between families, resulting in more person-to-person infections, compared to the populated core area with a larger component of households with one or two family members. Furthermore, there was a significant correlation in the total number of sentinel cases in the three seasons examined, which also suggests local characteristics increase the epidemic. Each house style did not seem to reflect such a community characteristic. Moreover, the age distribution of cases in these three seasons was consistent, with the highest ratio of infected children's groups being 0–4 years, followed by the 5–9 and 10–14 years age groups, respectively. These results are similar to those from another study [6] and suggest that the largest susceptibility to NoV exists in infants or

https://doi.org/10.1017/S0950268814000697 Published online by Cambridge University Press

younger children. However, our conclusions may also be due to a possibly higher frequency of younger children visiting clinics with minor symptoms.

A previous study suggested the importance of person-to-person transmission in the acquisition of NoV-associated infectious intestinal disease and it concluded that for young children (<5 years), contacts outside the household present a higher risk, whereas for older children (5–15 years) and adults, the highest risk is associated with child contacts inside the household [30]. These observations suggest that young children acquire their infections outside the home, and then put older children and adults at risk by contact at home. The peripheral area demographics suggest the incidence of such an event to be higher than for the central area.

The reason why this epidemic pattern recurred over the three seasons remains unclear. Large outbreaks often occur in closed settings, such as nursing homes, hospitals, and schools [21, 31]. The recurring pattern could suggest virus transmission in the peripheral area being related to these outbreak settings, although these data were not covered in the current study.

The 2006–2007 season had the highest recorded NoV epidemic in Japan because it was generated by the GII.4 2006b subtype. This virus variant remained predominant in the following 2007–2008 and 2008–2009 seasons [32]. The small decrease of incidence in the following 2007–2008 and 2008–2009 seasons may represent developed immunity to the virus variant after the largest epidemic in the 2006–2007 season, although such immunity may be limited to the short term [20]. NoV susceptibility remains [9, 21–23] and its incidence has stayed relatively high over the seasons. It is more likely to be local transmission which has affected the increasing number of cases unless there exists more susceptibility or any other factor in the peripheral area. Furthermore, the similar patterning over the seasons suggests a possible modelling of the epidemic as well as preventive measures with focus on the peripheral area should be sought. Thus a study of such factors would be important for considering the epidemic development.

It is not clear if the pattern of spread has resulted from the new GII.4 subtype. Moreover, a limitation of this study is that sentinel sites were mainly paediatric clinics, meaning data are primarily for children despite people of all ages being at risk for NoV. Although gastroenteritis data contain other causes such as rotavirus, the period of data used differs from the epidemic season of rotavirus which begins after the NoV season, about the end of February in Japan [25]. Therefore it is believed that our data consisted of mainly NoV data. Meanwhile each coefficient between the sentinel reported cases and the demographic data was calculated using a 1 km radius area. This radius is ideal, as patients commute from random distances and routes to each clinic. This study is also not supported by the background data of local medical standards. NoV epidemics occur with a complexity of factors that deserve a longer time-series study to identify other relevant indices such as the virus variant type and climate factors.

This is the first study to consider spreading patterns of NoV epidemics using sentinel surveillance data and demographic data.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268814000697.

ACKNOWLEDGEMENTS

This study was supported by a national research grant from the Ministry of Health, Labour, and Welfare, Japan. The sentinel data were obtained from NESID of NIID in Japan. The authors are grateful to Dr Kiyosu Taniguchi from the National Mie Hospital for his kind advice to this study.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Manish M, *et al.*** Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerging Infectious Diseases* 2008; **14**: 1224–1231.
2. **Hall AJ, *et al.*** Incidence of acute gastroenteritis and role of norovirus, Georgia, USA, 2004–2005. *Emerging Infectious Diseases* 2011; **17**: 1381–1388.
3. **de Wit MA, *et al.*** Gastroenteritis in sentinel general practices, The Netherlands. *Emerging Infectious Diseases* 2001; **7**: 82–91.
4. **Wheeler JG, *et al.*** Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *British Medical Journal* 1999; **318**: 1046–1050.
5. **Huhulescu S, *et al.*** Etiology of acute gastroenteritis in three sentinel general practices, Austria 2007. *Infection* 2009; **37**: 103–108.
6. **Bernard H, *et al.*** Epidemiology of norovirus gastroenteritis in Germany 2001–2009: eight seasons of

routine surveillance. *Epidemiology and Infection* 2013; **21**: 1–12.
7. **Mounts AW, *et al.*** Cold weather seasonality of gastroenteritis associated with Norwalk-like viruses. *Journal of Infectious Diseases* 2000; **181** (Suppl. 2): S284–287.
8. **Siebenga JJ, *et al.*** High prevalence of prolonged Norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution. *Journal of Infectious Diseases* 2008; **198**: 994–1001.
9. **Rockx B, *et al.*** Natural history of human calicivirus infection: a prospective cohort study. *Clinical Infectious Diseases* 2002; **35**: 246–253.
10. **Arvelo W, *et al.*** Norovirus outbreak of probable waterborne transmission with high attack rate in a Guatemalan resort. *Journal of Clinical Virology* 2012; **55**: 8–11.
11. **Centers for Disease Control and Prevention (CDC).** Recurring norovirus outbreaks in a long-term residential treatment facility – Oregon, 2007. *Morbidity and Mortality Weekly Report* 2009; **58**: 694–698.
12. **Kuusi M, *et al.*** A prolonged outbreak of Norwalk-like calicivirus (NLV) gastroenteritis in a rehabilitation centre due to environmental contamination. *Epidemiology and Infection* 2002; **129**: 133–138.
13. **Fankhauser RL, *et al.*** Molecular epidemiology of 'Norwalk-like viruses' in outbreaks of gastroenteritis in the United States. *Journal of Infectious Diseases* 1998; **178**: 1571–1578.
14. **Marks PJ, *et al.*** Evidence for airborne transmission of Norwalk-like virus (NLV) in a hotel restaurant. *Epidemiology and Infection* 2000; **124**: 481–487.
15. **Marks PJ, *et al.*** A school outbreak of Norwalk-like virus: evidence for airborne transmission. *Epidemiology and Infection* 2003; **131**: 727–736.
16. **Phillips G, *et al.*** Risk factors for symptomatic and asymptomatic norovirus infection in the community. *Epidemiology and Infection* 2011; **139**: 1676–1686.
17. **Dingle KE, *et al.*** Mutation in a Lordsdale norovirus epidemic strain as a potential indicator of transmission routes. *Journal of Clinical Microbiology* 2004; **42**: 3950–3957.
18. **Lindesmith L, *et al.*** Human susceptibility and resistance to Norwalk virus infection. *Nature Medicine* 2003; **9**: 548–553.
19. **Lindesmith LC, *et al.*** Mechanisms of GII.4 norovirus persistence in human populations. *PLoS Medicine* 2008; **5**: e31.
20. **Siebenga JJ, *et al.*** Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. *Journal of Virology* 2007; **81**: 9932–9941.
21. **Centers for Disease Control and Prevention (CDC).** Norovirus activity – United States, 2006–2007. *Morbidity and Mortality Weekly Report* 2007; **56**: 842–846.
22. **Barker J, *et al.*** Effects of cleaning and disinfection in reducing the spread of Norovirus contamination via environmental surfaces. *Journal of Hospital Infection* 2004; **58**: 42–49.
23. **Evans MR, *et al.*** An outbreak of viral gastroenteritis following environmental contamination at a concert hall. *Epidemiology and Infection* 2002; **129**: 355–360.
24. **Motomura K, *et al.*** Identification of monomorphic and divergent haplotypes in the 2006–2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history. *Journal of Virology* 2008; **82**: 11 247–11 262.
25. **Inaida S, *et al.*** The south to north variation of Norovirus epidemics from 2006–07 to 2008–09 in Japan. *PLoS ONE* 2013; **8**: e71 696.
26. **Taniguchi K, *et al.*** Overview of infectious disease surveillance system in Japan, 1999–2005. *Journal of Epidemiology* 2007; **17** (Suppl.): S3–13.
27. **Infectious Agents Surveillance Report (IASR).** Infectious Surveillance Centre, NIID, Japan (http://idsc.nih.go.jp/iasr/prompt/graph-ke.html).
28. **National Institute of Infectious Diseases (NIID).** Infectious gastroenteritis cases reported per sentinel weekly (http://www.nih.go.jp/niid/ja/10/2096-weeklygraph/1647-04gastro.html).
29. **Carrat F, *et al.*** Epidemiologic mapping using the 'kriging' method: application to an influenza-like illness epidemic in France. *American Journal of Epidemiology* 1992; **135**: 1293–1300.
30. **Phillips G, *et al.*** Risk factors for symptomatic and asymptomatic norovirus infection in the community. *Epidemiology and Infection* 2011; **139**: 1676–1686.
31. **Blanton LH, *et al.*** Molecular and epidemiologic trends of caliciviruses associated with outbreaks of acute gastroenteritis in the United States, 2000–2004. *Journal of Infectious Diseases* 2006; **193**: 413–421.
32. **Motomura K, *et al.*** Divergent evolution of Norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan. *Journal of Virology* 2010; **84**: 8085–8097.