

## Characteristics of child daycare centres associated with clustering of major enteropathogens

R. PIJNACKER<sup>1\*</sup>, L. MUGHINI-GRAS<sup>1,2</sup>, H. VENNEMA<sup>1</sup>, R. ENSERINK<sup>1</sup>,  
C. C. VAN DEN WIJNGAARD<sup>1</sup>, T. KORTBEEK<sup>1</sup> AND W. VAN PELT<sup>1</sup>

<sup>1</sup> Center for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), The Netherlands

<sup>2</sup> Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, The Netherlands.

Received 26 September 2015; Final revision 12 February 2016; Accepted 2 May 2016;  
first published online 25 May 2016

### SUMMARY

Insights into transmission dynamics of enteropathogens in children attending daycare are limited. Here we aimed at identifying daycare centre (DCC) characteristics associated with time-clustered occurrence of enteropathogens in DCC-attending children. For this purpose, we used the KIzSS network, which comprises 43 DCCs that participated in infectious disease surveillance in The Netherlands during February 2010–February 2013. Space–time scan statistics were used to identify clusters of rotavirus, norovirus, astrovirus, *Giardia lamblia* and *Cryptosporidium* spp. in a two-dimensional DCC characteristic space constructed using canonical correlation analysis. Logistic regression models were then used to further identify DCC characteristics associated with increased or decreased odds for clustering of enteropathogens. Factors associated with increased odds for enteropathogen clustering in DCCs were having indoor/outdoor paddling pools or sandpits, owning animals, high numbers of attending children, and reporting outbreaks to local health authorities. Factors associated with decreased odds for enteropathogen clustering in DCCs were cleaning child potties in designated waste disposal stations, cleaning vomit with chlorine-based products, daily cleaning of toys, extra cleaning of toys during a suspected outbreak, and excluding children with gastroenteritis. These factors provide targets for reducing the burden of gastrointestinal morbidity associated with time-clustered occurrence of major enteropathogens in DCC attendees.

**Key words:** Childhood infection, cluster, daycare, epidemiology, risk factor.

### INTRODUCTION

More than half of preschool children in The Netherlands (aged <4 years) experience at least one gastroenteritis (GE) episode each year [1]. About 40% of children in this age group are cared for in daycare centres (DCCs),

where they are at increased risk for GE and infectious diseases in general [2–4]. A recent Dutch study reported that these children are almost twice as likely to experience GE needing medical attention as expected based on national estimates [5]. This is comparable with other studies from the United States, reporting out-of-home childcare to be associated with a relative risk for diarrhoeal illness in preschool children of 1·8 compared to home care [6]. Although most of these GE episodes are mild and self-limiting, they may have an impact on parents/caregivers and on the society as a

\* Author for correspondence: Mr R. Pijnacker, National Institute for Public Health and the Environment (RIVM), Centre for Infectious Disease Control (CIb), PO Box 1–3720 BA Bilthoven, The Netherlands.  
(Email: roan.pijnacker@rivm.nl)

whole via productivity losses (e.g. days of work lost for care) and increased expenditure on healthcare services and treatment [1, 7]. Decreasing gastrointestinal morbidity in children attending DCCs requires insights into transmission dynamics and targets for control, which are currently not well established.

A recent Dutch study has identified several socio-demographic, facility- and policy-related DCC characteristics associated with GE incidence as a whole and with the prevalence of specific enteropathogens in DCC attendees [8]. However, this study did not distinguish between sporadic and clustered occurrence of these enteropathogens, while their risk factors may well differ. Space–time scan statistics have been applied to investigate clustering of infectious agents in hospitals and other community care services [9–12], and may also help identify clustering of enteropathogens in DCC-attending children and risk factors involved.

In this study, we exploit projections of DCCs in a two-dimensional space based on similarities of DCC characteristics, making it possible to apply space–time scan statistics to detect temporal clusters of pathogen detections in DCC attendees in relation to DCC characteristics. Clusters of pathogen occurrence are defined here as non-trivial groupings of individual pathogen detections as revealed by ‘expectation-based’ scan statistics. These statistics fit baseline models to the observed data in order to identify where and when there have been anomalous numbers of pathogen detections given those expected. Clusters can present themselves in different forms. They can have varying time spans and can either be restricted to small areas (i.e. associated with one or a few DCC characteristics) or be widespread (i.e. associated with many DCC characteristics). Outbreaks, seasonal and non-seasonal increases in pathogen circulation may all result in clustering. Although an increase in pathogen occurrence may mirror an increase in pathogen activity, it does not necessarily mean an excess of disease burden, as an infection can also be asymptomatic. While seasonality usually results in periodic (year-to-year) clustering, outbreaks and non-seasonal increases are more likely to manifest themselves irregularly. Investigating pathogen clusters using scan statistics may therefore embrace several potential outcomes, offering the opportunity of generating novel epidemiological insights towards understanding a pathogen’s temporal pattern and risk factors involved.

The aims of this study were to: (1) assess potential clustering of five major childhood enteropathogens (rotavirus, norovirus, astrovirus, *Giardia lamblia* and *Cryptosporidium* spp. [8]) in time as related to DCC

characteristics; (2) identify which socio-demographic, facility- and policy-related DCC characteristics were those associated with the clustered occurrence of these enteropathogens in DCC attendees; and (3) reveal possible differences with DCC characteristics associated with pathogen prevalence as a whole.

## MATERIALS AND METHODS

### Data sources

We used microbiological and epidemiological data from the KIzSS network, which comprises DCCs that participated in infectious disease surveillance in The Netherlands during 2010–2013 [13]. In March 2010, 2011 and 2012, all DCCs registered in the Dutch national database maintained by the Ministry of Social Affairs and Employment in The Netherlands were invited to participate in a dynamic DCC network for the surveillance of enteropathogens in their child populations. A detailed description of the DCC network design, methodology, definitions and limitations has been reported elsewhere [8, 13].

Forty-three DCCs agreed to submit 10 randomly selected faecal samples on a monthly basis from their attending children aged up to 48 months, regardless of the presence of gastrointestinal symptoms at the time of sampling. Of these, only 36 DCCs actually submitted samples (Fig. 1). Only one faecal sample per child within the same month was allowed to be submitted. All faecal samples were tested for 16 different enteropathogens using multiple, internally controlled quantitative PCRs (Table 1) [13]. For detailed information about the molecular detection methods, we refer to a previously published open-access paper [1]. This study focused on five major enteropathogens that were significantly associated with GE incidence in DCC attendees [14], accounting for 39% of GE morbidity; 11% by rotavirus, 10% by norovirus, 8% by *G. lamblia*, 7% by astrovirus, and 3% by *Cryptosporidium* spp. [14].

At the time of recruitment, DCCs were asked to complete a questionnaire on their characteristics, covering three main categories: socio-demographics, facility design and policies for hygiene and disease control, of which a detailed description is reported elsewhere [8]. This study builds upon and uses the same data of a previous study [8] to which we refer for more information. While the previous study looked at associations between DCC characteristics and enteropathogen prevalence as a whole, the present study looked at associations between DCC characteristics



**Fig. 1.** Geographical location of the 36 daycare centres, represented by their ID numbers, included in this study, The Netherlands. ■, Daycare centre.

and clustered occurrence of these enteropathogens. Thus, the two studies differ in the outcome of interest and in the approach used. For the purpose of this study, we focused on DCC characteristics that were found to be significantly associated with the prevalence of rotavirus, norovirus, astrovirus, *G. lamblia* or *Cryptosporidium* spp. in DCC attendees in the previous study (Table 2) [8].

### Reconstructing the 'spatial' distribution of DCC characteristics

Ideally, the true geographical locations of the DCCs would be used to identify space–time clusters of enteropathogens to infer physical transmission of enteropathogens in DCCs. However, since the 43 DCCs of this study were scattered all over The Netherlands

Table 1. *Enteropathogens included in the microbiological surveillance*

Virus	Bacteria	Parasites
Norovirus*	<i>Escherichia coli</i> ‡	<i>Giardia lamblia</i> *
Sapovirus	<i>Salmonella enterica</i>	<i>Cryptosporidium</i> spp.*
Rotavirus*	Shigella spp.	<i>Dientamoeba fragilis</i>
Adenovirus†	<i>Campylobacter jejuni</i>	
Astrovirus*	<i>Clostridium difficile</i> <i>Yersinia enterocolitica</i>	

\* The current study focused on these five enteropathogens.

† Enteropathogenic types 30 and 40/41.

‡ Shiga toxin producing *Escherichia coli* (STEC), enteroaggregative *E. coli* (EAEC) and typical and atypical enteropathogenic *E. coli* (EPEC).

(Fig. 1), any cluster incorporating multiple DCCs was unlikely to be the result of direct enteropathogens transmission between DCCs, as they were located too far apart from each other to be physically related to one another. Therefore, we looked at temporal clustering of enteropathogens over infrastructural similarities, rather than geographical proximities of the DCCs. To this end, we constructed an alternative spatial distribution for the participating DCCs using canonical correlation analysis (CCA). CCA is a statistical method that finds the multivariate relation(s) between two sets of variables (DCC characteristics in one set and enteropathogens in the other), with each set consisting of two or more variables [15]. CCA was used here to calculate linear combinations of DCC characteristics (so called canonical variates) that maximize relationships to a similarly calculated linear combination of enteropathogens. This means that DCCs were positioned in the reconstructed space according to their groupings of common characteristics, with those DCCs positioned closer to each other in the reconstructed space being those more similar to one another with regard to their characteristics putatively associated with enteropathogen occurrence, with the opposite applying for those DCCs that were placed far apart in the reconstructed space (Fig. 2). This alternative space, where the DCCs were placed, was defined by the first and second canonical variates, hereafter referred to as DCC characteristic space (Fig. 3).

### Cluster analysis

The series of faecal samples submitted between February 2010 and February 2013 were analysed retrospectively for the presence of rotavirus,

norovirus, astrovirus, *G. lamblia* and *Cryptosporidium* spp. temporal clusters in the DCC characteristic space. Analyses were performed using month-level Bernoulli models in SaTScan v. 9.1.1 [16]. The maximum temporal window for a cluster was set to 50% of the study period (18 months) and a maximum of 50% of the study population was allowed to be included in a cluster. The scan statistic is defined by a cylindrical window varying in size, with a circular spatial base and with height corresponding to time, which evaluates each possible time period, centring the window on each DCC in the DCC characteristic space. For each DCC, only the cylinder with the maximum likelihood is reported as the most likely cluster, which is calculated based on the number of observed and expected number of pathogens inside and outside the cylinder. A *P* value is assigned via 999 Monte Carlo simulations [17]. Scan statistics were performed on six temporally contiguous time windows of 6 months each, from February 2010 to January 2013. In addition, six additional temporally contiguous time windows were analysed with a 6-month interval, from May 2010 to February 2013. The clusters so identified, represented increased occurrences in time of enteropathogen-positive faecal samples in children attending DCCs with similar characteristics.

### Analysis of DCC characteristics associated with clustering

Clusters identified using SaTScan with  $P > 0.05$  by the likelihood ratio test were discarded. Overlapping clusters in time and DCC were merged for each pathogen. DCCs were excluded from a cluster when/if they did not send any samples during the time span of the cluster. Based on the results of the cluster analyses, a dataset was constructed in which each DCC was either included in a cluster (encoded as 1) or not (encoded as 0) at each month of the study period, thereby obtaining a binary dataset of 1332 DCC-month observations (i.e. 36 DCCs followed-up for 37 months) for each pathogen. Pathogen presence or absence was assigned to a month based on the first day of the week the faecal sample was collected.

We then assessed whether the aforementioned 21 DCC characteristics were significantly associated with a higher or lower odds for a DCC to be included in a space-time cluster (i.e. those identified using SaTScan) of rotavirus, norovirus, astrovirus, *G. lamblia* or *Cryptosporidium* spp. using logistic regression. A separate analysis was performed for each pathogen.

Table 2. Daycare centre (DCC) characteristics focussed on in this study that were significantly associated with the prevalence of rotavirus, norovirus, astrovirus, *G. lamblia* or *Cryptosporidium spp.* in DCC attendees in a previous study based on the KIzSS network [8]

Hygiene policies Cleaning vomit with paper* Cleaning vomit with chlorine-based products Cleaning toys more or less frequently than daily Cleaning bed linen more or less frequently than daily Cleaning child potties in a normal sink† Facility design Presence of dedicated toilets for ill children Presence of animals Presence of paddling pools Presence of sandpits Staff members have multiple daily duties	Control measures during suspected outbreak Excluding children with gastroenteritis Excluding personnel with gastroenteritis No mixing of staff members in child groups Extra cleaning of toys Notification to local health authorities Demographics Weekly number of attending children Mixing of personnel in child groups
--	---

\* Rather than cloth towels but without cleaning solution.

† Rather than in a designated waste disposal station.

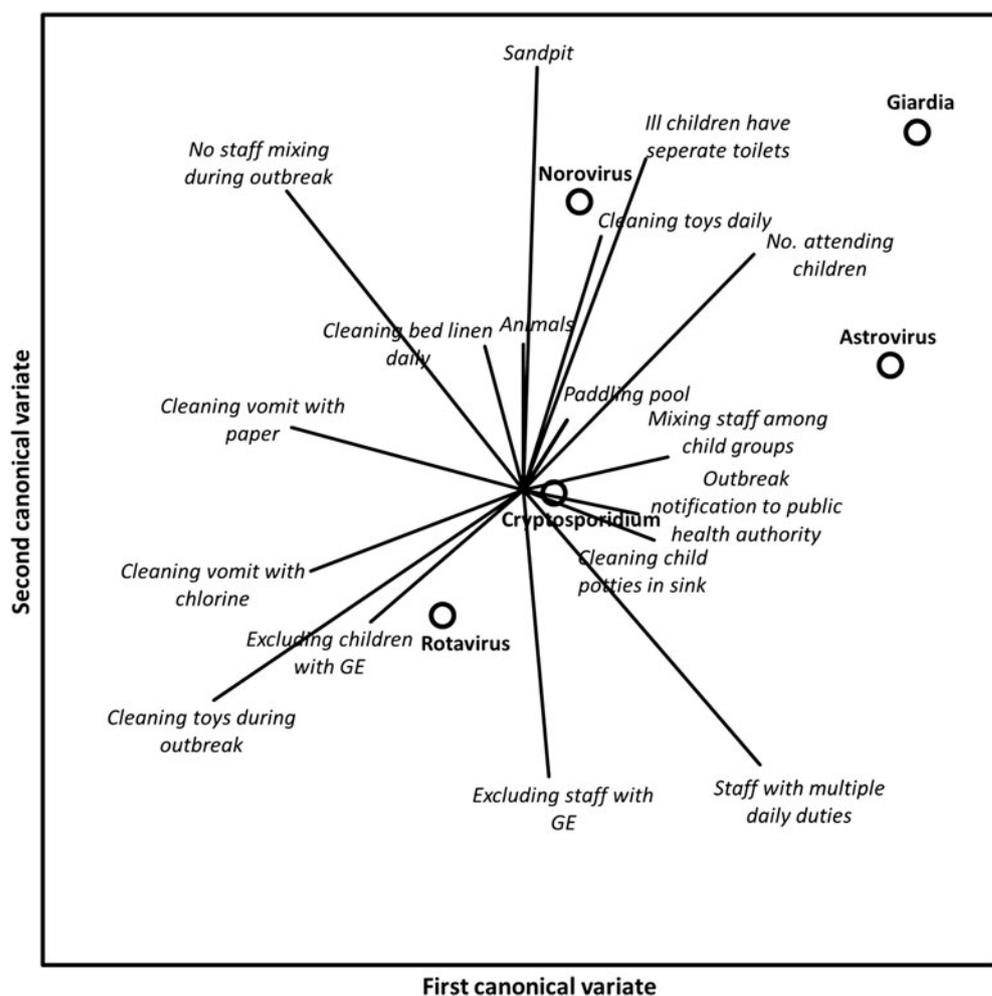
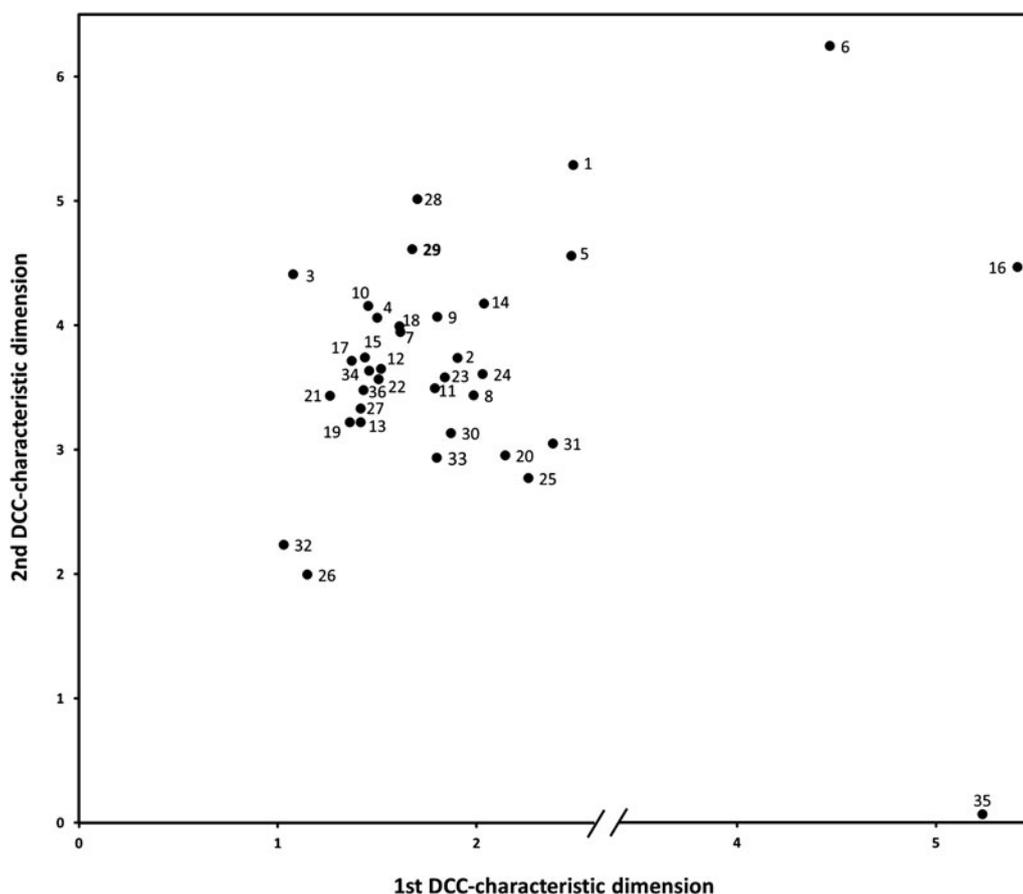


Fig. 2. Plot of the daycare centre (DCC) characteristics and enteropathogens over the first and second canonical variate, i.e. the two main dimensions of the DCC characteristic space (cf. Fig. 3). GE, Gastroenteritis.



**Fig. 3.** Reconstructed spatial distribution in the daycare centre (DCC) characteristic space of the 36 DCCs included in this study, represented by their ID numbers.

Associations of the 17 DCC characteristics listed in Table 2 were first assessed one by one in a baseline model including socioeconomic status (expressed as a normalized score from  $-4$  to  $+4$  based on income, employment and educational level per postal code area, with a higher score indicating lower socioeconomic status), urbanization degree [categorized as ‘highly urbanized’ ( $>2500$  addresses/km<sup>2</sup>), ‘urbanized’ (1500–2500 addresses/km<sup>2</sup>), ‘moderately urbanized’ (1000–1500 addresses/km<sup>2</sup>), ‘lowly urbanized’ (500–1000 addresses/km<sup>2</sup>), and ‘rural’ ( $<500$  addresses/km<sup>2</sup>)], season [categorized as ‘winter’ (December–February); ‘spring’ (March–May); ‘summer’ (June–August); ‘autumn’ (September–November)] and calendar year (2010 to 2013) as control variables. DCC characteristics with  $P \leq 0.10$  at the single-variable analysis were then entered in a multivariable logistic regression model built in backward stepwise fashion: only variables with  $P < 0.05$  were retained. The effect of removing variables from the model on the coefficients of other variables was also monitored. Collinearities between independent variables were

checked prior to regression analysis with cross-tabulations and collinear variables were selected based on improved model fit (Akaike’s Information Criterion) or combined in one variable when they predicted each other perfectly. Multivariable logistic regression models with an overall statistical significance (likelihood ratio  $\chi^2$  test,  $P < 0.05$ ) and goodness-of-fit (Hosmer–Lemeshow test,  $P > 0.05$ ) are reported. Seasonal and non-seasonal clusters were distinguished in order to identify whether risk factors might differ. Clusters occurring in the same season for three subsequent years were classified as seasonal, and non-seasonal otherwise. Statistical analyses were performed using Stata v. 13.0 software (StataCorp., USA).

## RESULTS

From February 2010 to February 2013, 5125 faecal samples were submitted by the 36 enrolled DCCs. Of these samples, 148 (2.9%) tested positive for

astrovirus, 478 (9.3%) for norovirus, 153 (3.0%) for rotavirus, 39 (0.8%) for *Cryptosporidium* spp. and 219 (4.3%) for *G. lamblia*.

### Clustering of enteropathogens

A total of 229 clusters over all five enteropathogens were identified during the study period (Fig. 4). For rotavirus, 79 clusters were found in 35 different DCCs (97% of all DCCs participating), including 82.4% of all samples found positive for rotavirus, with a time span of 1–5 months (median 2 months). For norovirus, 76 clusters were found in 27 DCCs (75%), including 77.8% of all samples found positive for norovirus, with a time span of 1–12 months (median 2 months). For astrovirus, 43 clusters were found in 31 DCCs (86%), including 70.9% of all samples found positive for astrovirus, with a time span of 1–4 months (median 2 months). For *G. lamblia*, 29 clusters were found in 15 DCCs (42%), including 53% of all samples found positive for *G. lamblia*, with a time span of 1–11 months (median 3 months). For *Cryptosporidium* spp., 22 clusters were found in 18 DCCs (50%), including 79.5% of samples found positive for *C. hominis*, with a time span of 1–5 months (median 2 months).

All clusters of rotavirus and astrovirus were classifiable as seasonal, whereas those of norovirus, *G. lamblia*, and *Cryptosporidium* spp. were classifiable as non-seasonal (Fig. 4). Therefore, we did not stratify our risk factor analysis by seasonal and non-seasonal clustering since all clusters for a given pathogen appeared to be of the same type. Seasonal clusters were observed during the whole study period for rotavirus, and from 2011 to 2013 for astrovirus. For norovirus, *G. lamblia* and *Cryptosporidium* spp. non-seasonal clusters were identified throughout the year. No clusters were found after January 2012 for norovirus, while clustered *Cryptosporidium* spp. occurrence was observed only in 2012. *G. lamblia* clusters were found throughout the study period, with a longer time span and fewer DCCs involved, compared to the other enteropathogens.

### DCC characteristics associated with clustering in time of *G. lamblia*

As shown in Table 3, five DCC characteristics were significantly associated with clustering of *G. lamblia* in DCCs. Keeping animals and having sandpits in the DCC was associated with an increased risk for DCCs to experience clustered occurrence of *G. lamblia* in their child populations [odds ratio (OR) 17.0, 95%

confidence interval (CI) 1.7–169.4]. Reporting any suspected GE outbreak to local health authorities was also associated with a higher risk of experiencing a cluster of *G. lamblia* occurrence (OR 4.1, 95% CI 1.7–9.5), so was the presence of paddling pools (OR 2.0, 95% CI 1.2–3.6). Cleaning vomit with chlorine-based products was associated with a decreased risk of DCCs having a *G. lamblia* cluster (OR 0.3, 95% CI 0.1–0.8). The same was found for extra cleaning of toys during a suspected GE outbreak (OR 0.3, 95% CI 0.1–0.6).

### DCC characteristics associated with clustering in time of *Cryptosporidium* spp.

Two DCC characteristics were associated with clustering of *Cryptosporidium* spp. in DCCs. Cleaning toys daily compared to cleaning toys less than daily and excluding children with GE from entering the DCC were both associated with a decreased risk for DCCs to experience clustered occurrence of *Cryptosporidium* spp. in their attendees (OR 0.3, 95% CI 0.1–0.9 and OR 0.1, 95% CI 0.0–0.4, respectively).

### DCC characteristics associated with clustering in time of rotavirus and astrovirus

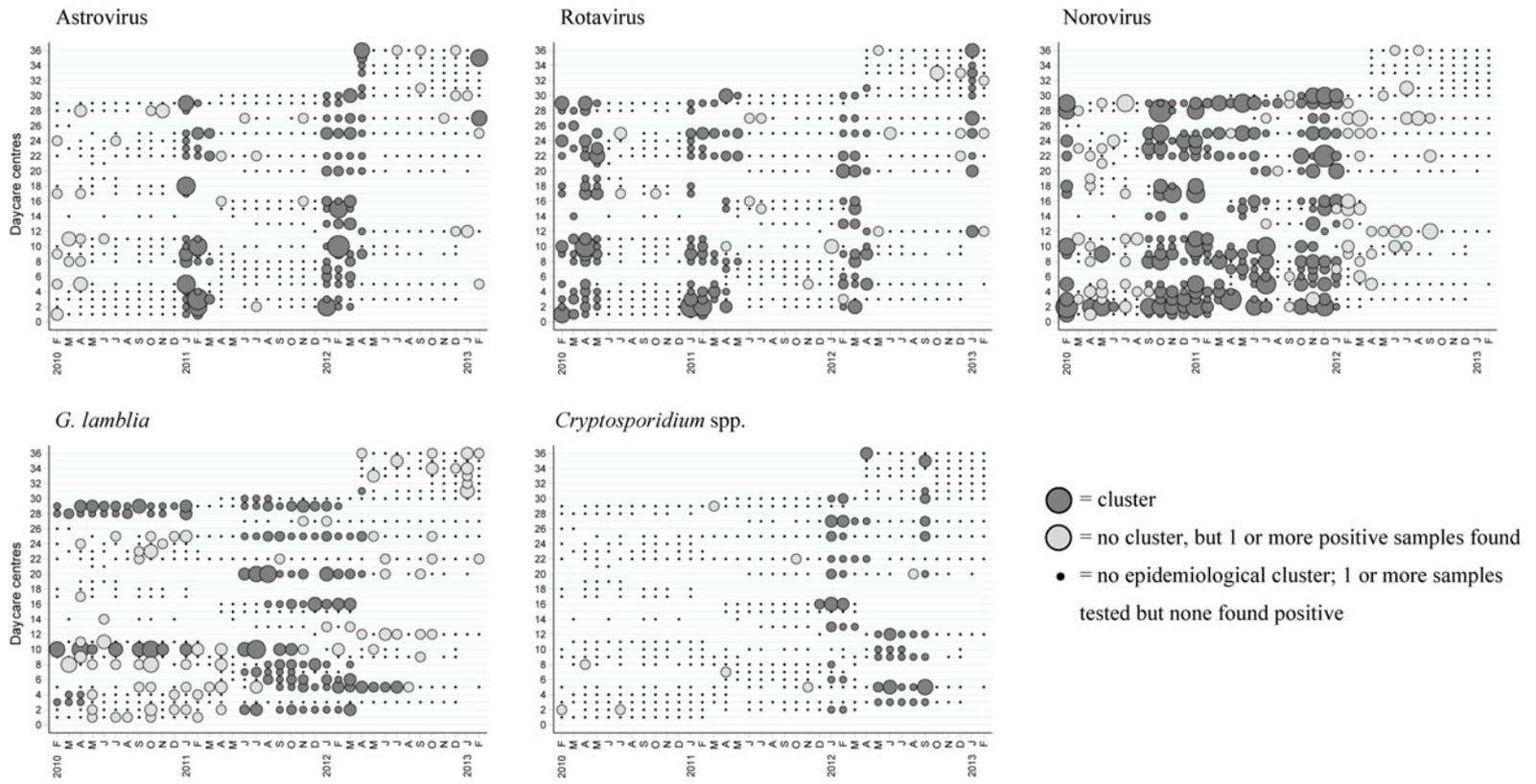
Only the presence of sandpits in the DCC was found to be associated with clustering of rotavirus in DCCs (OR 4.0, 95% CI 1.2–13.1). The same association was found for astrovirus (OR 3.9, 95% CI 1.2–13.2).

### DCC characteristics associated with clustering in time of norovirus

Five DCC characteristics were associated with clustering of norovirus in DCCs. A higher number of DCC attendees was associated with increased risk for DCCs to experience a cluster of norovirus occurrence (OR 2.4, 95% CI 1.6–3.6). DCCs with sandpits or paddling pool were at increased risk of experiencing clustered occurrence of norovirus in their attendees (OR 5.5, 95% CI 2.7–11.3 and OR 4.3, 95% CI 2.3–8.2, respectively). Cleaning child potties in a dedicated sink was associated with decreased risk of having norovirus clusters (OR 0.4, 95% CI 0.2–0.7), as was cleaning of toys during a suspected outbreak (OR 0.5, 95% CI 0.3–0.9).

## DISCUSSION

This study identified several DCC characteristics that were associated with temporal clustering of rotavirus,



**Fig. 4.** Clusters of astrovirus, rotavirus, norovirus, *G. lamblia* and *Cryptosporidium* spp. in Dutch daycare centres from February 2010 to February 2013. The size of the circle correlates with the number of faecal samples positive for the relevant pathogen; no circle indicates absence of samples submitted.

**Table 3.** Adjusted odds ratios and corresponding 95% confidence intervals for the association between daycare centre (DCC) characteristics and clustering of rotavirus, norovirus, astrovirus, *G. lamblia* and *Cryptosporidium* spp. in DCCs

	<i>G. lamblia</i> †	<i>Cryptosporidium</i> †	Rotavirus†	Astrovirus†	Norovirus†
<b>Demographics</b>					
DCC capacity (no. attending children)	n.s.	n.s.	n.s.	n.s.	2.40 (1.61–3.56)*
Socioeconomic status neighbourhood	2.34 (1.56–3.49)	0.64 (0.44–0.94)	0.79 (0.58–1.07)	0.89 (0.61–1.28)	1.10 (0.86–1.41)
Urbanization degree (addresses/km <sup>2</sup> )					
<500	Ref.	Ref.	Ref.	Ref.	Ref.
500–1000	0.69 (0.18–2.63)	0.28 (0.09–0.95)	1.31 (0.50–3.38)	2.01 (0.64–6.30)	4.78 (1.92–11.91)
1000–1500	0.92 (0.34–2.44)	0.11 (0.03–0.44)	0.80 (0.31–2.07)	1.35 (0.44–4.12)	3.43 (1.51–7.79)
1500–2500	1.18 (0.34–4.03)	0.08 (0.02–0.39)	0.77 (0.24–2.48)	1.33 (0.31–5.71)	4.56 (1.69–12.28)
>2500	6.46 (1.51–27.68)	0.20 (0.05–0.75)	1.32 (0.40–4.38)	1.81 (0.45–7.24)	2.14 (0.82–5.59)
<b>Facility design</b>					
Staff members with multiple daily duties	n.s.	n.s.*	n.s.	n.s.	n.s.
Mixing of staff members in child groups	n.s.	n.s.	n.s.*	n.s.	n.s.
Cleaning child potties in a designated waste disposal station	n.s.	n.s.	n.s.*	n.s.	0.37 (0.20–0.69)
Having indoor/outdoor sandpits	n.a.*	n.s.	4.01 (1.23–13.08)	3.90 (1.15–13.23)	5.50 (2.68–11.28)*
Having indoor/outdoor paddling pools	2.03 (1.16–3.55)	n.s.	n.s.	n.s.	4.30 (2.27–8.15)*
Owning animals (and having sandpits‡)	16.98 (1.70–169.40)*	n.s.	n.s.	n.s.	n.s.
<b>Hygiene policy</b>					
Cleaning vomit with chlorine-based products	0.29 (0.11–0.76) *	n.s.	n.s.	n.s.	n.s.
Cleaning vomit with paper towels	n.s.	n.s.	n.s.	n.s.	n.s.*
Daily cleaning of bed linen	n.s.	n.s.	n.s.	n.s.	n.s.*
Daily cleaning of toys	n.s.	0.27 (0.08–0.92)	n.s.	n.s.*	n.s.
Extra cleaning of toys during outbreak	0.25 (0.11–0.60)*	n.s.	n.s.	n.s.	0.54 (0.31–0.94)
Excluding children with gastrointestinal symptoms	n.s.	0.09 (0.02–0.37)	n.s.*	n.s.	n.s.
Excluding personnel with gastrointestinal symptoms	n.s.*	n.s.	n.s.	n.s.	n.s.
Ill children have separate toilets	n.s.	n.s.	n.s.*	n.s.	n.s.
Always reporting outbreaks to local health authority	4.05 (1.72–9.54)	n.s.	n.s.	n.s.*	n.s.
No mixing of staff members during outbreak	n.s.	n.s.	n.s.	n.s.	n.s.*
<b>Year</b>					
2010	Ref.	n.c.§	Ref.	n.c.	n.s.
2011	2.64 (1.46–4.78)		0.64 (0.30–1.37)	Ref.	
2012	1.36 (0.69–2.68)		0.18 (0.08–0.39)	3.30 (1.60–6.79)	
2013	n.c.		0.72 (0.21–2.46)	0.11 (0.02–0.54)	
<b>Season</b>					
Winter	Ref.	Ref.	Ref.	Ref.	Ref.
Spring	0.77 (0.37–1.57)	0.62 (0.28–1.38)	2.59 (1.44–4.66)	0.26 (0.13–0.51)	0.20 (0.11–0.36)
Summer	1.07 (0.55–2.11)	0.78 (0.37–1.67)	n.c.	n.c.	0.15 (0.08–0.28)
Autumn	1.28 (0.66–2.49)	0.45 (0.20–1.03)	n.c.	n.c.	0.78 (0.47–1.31)

n.s., Non-significant ( $P > 0.05$ ); n.a., non-assessable (perfect predictor); n.c. non-calculable (no clusters found).

† Adjusted for season, urbanization degree, socioeconomic status and year.

‡ Owning animals perfectly predicts having sandpits.

§ Excluded from the model since no clusters were identified in 2010 and 2013, and only one cluster was identified in 2011.

|| Goodness-of-fit of the model improved with year excluded.

\* Significantly associated with pathogen prevalence as a whole in a previous study based on the same KIzSS network [8].

norovirus, astrovirus, *G. lamblia* and *Cryptosporidium* spp. in DCCs. Using the same data from the KIZSS network, differences were found between DCC characteristics associated with the clustered occurrence of these enteropathogens and those previously associated with pathogen prevalence as a whole [8]. This is the first study in which a combination of different methods like multivariate, clustering and regression analyses was used to define the epidemiological pattern of clusters of these enteropathogens in DCCs and the associated risk factors.

We found enteropathogen clusters to be associated with having sandpits and paddling pools in the DCC. Specifically, DCCs with sandpits were at increased risk of experiencing clusters of rotavirus, astrovirus, norovirus and *G. lamblia*. In the previous study [8], only norovirus and *G. lamblia* were associated with increased (albeit not necessarily clustered) prevalence of norovirus and *G. lamblia* in DCC attendees in DCCs with sandpits [8]. This might suggest that DCCs with sandpits are significantly more likely to experience clustering of rotavirus and astrovirus rather than increased prevalence of these pathogens *per se*. Contamination of sandpits with faecal coliforms has been reported in a Japanese study in public park sandpits [18] and a documented outbreak of norovirus occurred in a nursery school as a result of a child vomiting in a sandpit [19]. For the clustered occurrence of rotavirus and astrovirus, having a sandpit was the only risk factor identified, reflecting the similarities in the epidemiology of these two viruses. Paddling pools were also associated with increased occurrence of *G. lamblia* and norovirus clusters. As all DCCs had an outdoor play area, sandpits and paddling pools as risk factors for enteropathogen clustering are unlikely to be proxies for outdoor play. In the previous study [8], paddling pools were associated with increased prevalence of norovirus, but not with that of *G. lamblia*. Similar to sandpits, it can be speculated that the presence of paddling pools would result in increased clustering of *G. lamblia*. *G. lamblia* is known for its ability to be transmitted through contaminated water [20], and the vulnerability of swimming pools to norovirus contamination – even in the absence of any obvious vomiting or ‘faecal accident’ has been claimed [21].

Several hygiene-related DCC characteristics were also associated with clustering of enteropathogens. This is in agreement with previous studies who showed that disinfection of inanimate objects may be important in controlling the spread of enteric diseases in

DCCs [22, 23]. A decreased prevalence of astrovirus was observed in DCCs that cleaned toys daily in the previous study [8]. Additionally, a decreased prevalence of *G. lamblia* was previously observed in DCCs performing extra cleaning of toys during a suspected outbreak [8], a measure that also decreased the clustered occurrence of *G. lamblia* in this study. However, cleaning toys daily did not decrease the occurrence of astrovirus clusters. Decreased clustered occurrence of norovirus and *Cryptosporidium* spp. was observed in DCCs performing extra cleaning of toys during outbreaks and in those cleaning toys daily, respectively. These findings suggest that the effectiveness of disinfection of inanimate objects might depend on the nature of enteropathogen occurrence, which can be sporadic or clustered. *G. lamblia* has been recovered from chairs and tables in DCCs while norovirus is known to persist for days on dry inanimate surfaces [24]. Cleaning vomit with chlorine-based products was previously associated with decreased prevalence of *G. lamblia* [8], and the same association was found here with clustering. This may reflect the effectiveness of chlorine in inactivating this protozoan parasite [25]. Clustering of norovirus occurred significantly less often in DCCs that cleaned child potties in a designated waste disposal station rather than in a normal sink, and a high number of attending children was associated with increased clustering of norovirus, reflecting its potential for person-to-person transmission [26]. A high number of attending children was also previously associated with increased norovirus prevalence, while cleaning child potties in a designated waste disposal station (rather than in a normal sink) was not [8]. This suggests that cleaning child potties in a designated station might prevent norovirus from occurring as clusters, but not from occurring as a whole.

In agreement with evidence provided for the value of exclusion policies for children with gastrointestinal symptoms in childcare facilities [27], we found that excluding children with GE significantly decreased the occurrence of *Cryptosporidium* spp. clusters. This policy was only associated with decreased prevalence of rotavirus in the previous study [8]. Interestingly, the latter study found that factors entailing transmission of enteropathogens via staff members, e.g. allowing them to have multiple daily duties, mixing with child groups, and entering the DCC only if/when free of gastrointestinal symptoms, were all associated with increased prevalence of one or more of the pathogens. However, none of these factors was associated with clustered occurrence of these pathogens. A possible

explanation is that transmission via staff members does occur, but it does not result in significant clusters.

Generally, DCC characteristics associated with *G. lamblia* clustering found here were similar to those previously associated with its prevalence [8]. A peculiar risk factor for *G. lamblia* clustering was whether DCCs always reported outbreaks to local health authorities, indicating that DCCs are more likely to always report to the local health authority if they have (had) problems in controlling infection themselves. In the previous study [8], this DCC policy was not associated with *G. lamblia*, but only with astrovirus.

All clusters of rotavirus and astrovirus were classifiable as seasonal, and those of norovirus, *G. lamblia*, and *Cryptosporidium* spp. were all classifiable as non-seasonal. Astrovirus clusters peaked during winter months, which is in line with previous studies [8, 14]. In agreement with a recent study in The Netherlands and findings in Japan [28, 29], we found rotavirus clusters mainly during spring. Clusters of *Cryptosporidium* spp. were only observed in 2012, which is in accordance with the documented increase in *C. hominis* infections in the late summer of 2012 in patients presenting to their general practitioner (GP) in The Netherlands, UK and Germany [30]. Remarkably, clusters in DCCs occurred early in the year, starting in January, while the increase in GP patients was observed in July [30]. This suggests that the (predominantly asymptomatic) increase in clustering of *Cryptosporidium* spp. in DCCs might have been an early indication for the corresponding increase of symptomatic cases in the general population. *G. lamblia* had the highest median time span of clusters compared to the other enteropathogens, a reflection of its relatively continuous occurrence throughout the study period [31], thereby not showing seasonality. Surprisingly, no seasonality was observed for norovirus clusters. The occurrence of either outbreaks or sporadic cases of norovirus is often characterized by a strong seasonality, with the majority of them occurring during the winter [32]. A study conducted in England and Wales [33] found that norovirus outbreaks in community settings like hospitals and residential facilities were more common from November to April than during the rest of the year. However, they also found that outbreaks in settings like private homes, holiday camps, and military bases displayed no winter peak. Although the latter study only focused on outbreaks, this would provide an explanation as to why there seems to be no clear seasonal pattern in the occurrence of norovirus clusters. However, our stool

samples were taken from both symptomatic and asymptomatic children. This sampling scheme might detect norovirus also when exposure is low and infection has a higher probability to evolve asymptotically, resulting in a non-seasonal pattern. In 2012 and 2013, no clusters of norovirus were identified, which may be explained by the way the SaTScan algorithm calculates the number of expected cases based on those occurring during the whole study period. Indeed, the relatively low prevalence of norovirus during 2012 and 2013 resulted in the absence of clusters detected in these years, with the exception of January 2012.

This study has several limitations. The questionnaires were self-administered by the DCCs. Although it is unlikely that the reported facility characteristics were not reported truthfully, it is possible that there were discrepancies between the actual and perceived behaviours towards policies in place in the DCC. Another limitation is that we could not differentiate between actual preventive measures taken against suspected outbreaks or only intentions to do so, or if this has been changed over time. This is because the data did not allow us to identify DCCs suspecting an outbreak during the study period and how they handled that. We were also unable to differentiate between enteropathogens that were introduced, but not transmitted in the DCC environment, and enteropathogens transmitted within the DCC. Yet, it seems unlikely that the introduction of an enteropathogen via one DCC attendee would lead to clustering if not further transmitted within the DCC. This is because the identified clusters were always comprised of multiple faecal samples positive for the enteropathogen in question, meaning that multiple DCC attendees would have introduced the same enteropathogen in a short time period, which is less likely to happen. Selection bias may also have occurred because faecal samples were only obtained from children while they were attending the DCC and not from children who were absent due to illness. However, pathogen shedding can also occur before and after the symptomatic phase, which is likely to have been detected by our sampling scheme. Since none of the enteropathogens of this study were bacteria, they were unlikely to have been eradicated by antibiotic treatments and they may still have been shed after the child re-entered the DCC. Besides, antibiotic treatment in The Netherlands is prescribed in <1% of children with GE visiting the GP [5]. Although DCCs were asked to sample children at random, the possibility that some samples were not collected completely at

random cannot be excluded. Finally, in the risk factor analysis, clusters were treated as present or absent at any point in time (month) independently of their size (i.e. number of samples positive for the pathogen in question). However, such data structure took into account the cluster duration, which is correlated to cluster size. Future improvements in this field should consider differentiating between clusters based on their size.

## CONCLUSIONS

This study provides novel insights in the socio-demographic, facility- and policy-related DCC characteristics associated with enteropathogen clustering in DCC attendees. It also indicates that consideration of the nature of enteropathogen occurrence is important when determining targets for control, as risk factors for clustering may be different than those for enteropathogen prevalence as a whole. This study provides targets for reducing the burden of gastrointestinal morbidity in DCC-attending children, such as sandpits and paddling pools, which were the main risk factor identified for the clustered occurrence of several enteropathogens. It also highlights the importance of hygiene policies enforced in the DCC, as cleaning vomit with chlorine-based products, cleaning toys (during suspected outbreaks), and cleaning child potties in designated waste disposal stations were found to reduce the risk of experiencing enteropathogen clustering. These findings help to understand how DCC characteristics can be related to the observed clustering pattern and may enhance future interventions.

## ACKNOWLEDGEMENTS

This study was conducted according to the principles of the Declaration of Helsinki. The Dutch Central Committee on Research involving Human Subjects in Utrecht, The Netherlands, gave ethical approval to conduct this study. No subject-identifiable results are generated. This paper presents independent research commissioned by the Dutch National Institute for Public Health and the Environment (RIVM) and has been internally financed.

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **Enserink R, et al.** High detection rates of enteropathogens in asymptomatic children attending day care. *PLoS ONE* 2014; **9**: e89496.
2. **Barros AJ.** Child-care attendance and common morbidity: evidence of association in the literature and questions of design. *Revista de Saude Publica* 1999; **33**: 98–106.
3. **Nesti MM, Goldbaum M.** Infectious diseases and day-care and preschool education. *Jornal de Pediatria* 2007; **83**: 299–312.
4. **Central Bureau for Statistics.** Less children attending day-care [in Dutch]. (<http://statline.cbs.nl>).
5. **Enserink R, et al.** Infectious disease burden related to child day care in the Netherlands. *Pediatric Infectious Disease Journal* 2013; **32**: e334–340.
6. **Lu N, et al.** Child day care risks of common infectious diseases revisited. *Child: Care, Health & Development* 20014; **30**: 8.
7. **Frieseema IH, et al.** Costs of gastroenteritis in the Netherlands, with special attention for severe cases. *European Journal of Clinical Microbiology & Infectious Diseases* 2012; **31**: 1895–1900.
8. **Enserink R, et al.** Risk factors for gastroenteritis in child day care. *Epidemiology and Infection* 2015; **143**: 2707–2720.
9. **Doyle TJ, et al.** Outbreaks of noroviral gastroenteritis in Florida, 2006–2007. *Epidemiology and Infection* 2009; **137**: 617–625.
10. **Jones RC, et al.** Use of a prospective space-time scan statistic. *Public Health Reports* 2006; **121**: 7.
11. **Horst MA, Coco AS.** Observing the spread of common illnesses through a community: using Geographic Information Systems (GIS) for surveillance. *Journal of the American Board of Family Medicine* 2010; **23**: 32–41.
12. **Davis GS, Sevdalis N, Drumright LN.** Spatial and temporal analyses to investigate infectious disease transmission within healthcare settings. *Journal of Hospital Infection* 2014; **86**: 227–243.
13. **Enserink R, et al.** The KIzSS network, a sentinel surveillance system for infectious diseases in day care centers: study protocol. *BMC Infectious Diseases* 2012; **12**: 259.
14. **Enserink R, et al.** Gastroenteritis attributable to 16 enteropathogens in children attending day care. Significant effects of rotavirus, norovirus, astrovirus, *Cryptosporidium* and *Giardia*. *Pediatric Infectious Disease Journal* 2014; **34**: 5–10.
15. **ter Braak CJF.** Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 1986; **67**: 12.
16. **Kulldorff M, Nagarwalla N.** Spatial disease clusters: detection and interference. *Statistics in Medicine*. 1995; **14**: 11.
17. **Kulldorff M.** A spatial scan statistic. *Communications in Statistics – Theory and Methods* 1997; **26**: 15.
18. **Matsuo J, Nakashio S.** Prevalence of fecal contamination in sandpits in public parks in Sapporo City, Japan. *Veterinary Parasitology* 2005; **128**: 115–119.
19. **O'Neill HJ, et al.** Gastroenteritis outbreaks associated with Norwalk-like viruses and their investigation by nested RT-PCR. *BMC Microbiology* 2001; **1**.

20. **Baldursson S, Karanis P.** Waterborne transmission of protozoan parasites: review of worldwide outbreaks – an update 2004–2010. *Water Research* 2011; **45**: 6603–6614.
21. **Podewils LJ, et al.** Outbreak of norovirus illness associated with a swimming pool. *Epidemiology and Infection* 2007; **135**: 827–833.
22. **Weniger BG, et al.** Fecal coliforms on environmental surfaces in two day care centers. *Applied and Environmental Microbiology* 1983; **45**: 3.
23. **European Centre for Disease Prevention and Control.** Prevention of norovirus infection in schools and child-care facilities, 2013.
24. **Kramer A, Schwebke I, Kampf G.** How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infectious Diseases* 2006; **6**: 130.
25. **Lyman WH, et al.** Prospective study of etiologic agents of acute gastroenteritis outbreaks in child care centers. *Journal of Pediatrics* 2009; **154**: 5.
26. **European Centre for Disease Prevention and Control.** **Annual epidemiological report 2013.** Reporting on 2011 surveillance data and 2012 epidemic intelligence data. Stockholm: ECDC, 2013.
27. **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.
28. **Suzuki H, et al.** Peak rotavirus activity shifted from winter to early spring in Japan. *Pediatric Infectious Disease Journal* 2005; **24**: 257–260.
29. **Hahne S, et al.** Exceptionally low rotavirus incidence in the Netherlands in 2013/14 in the absence of rotavirus vaccination. *Eurosurveillance* 2014; **19**.
30. **Fournet N, et al.** Simultaneous increase of Cryptosporidium infections in the Netherlands, the United Kingdom and Germany in late summer, 2012. *Eurosurveillance*. 2013; **18**.
31. **Enserink R, et al.** High detection rates of enteropathogens in asymptomatic children attending day care. *PLoS ONE* 2014; **9**: 7.
32. **Ahmed SM, Lopman BA, Levy K.** A systematic review and meta-analysis of the global seasonality of norovirus. *PLoS ONE* 2013; **8** : e75922.
33. **Lopman BA, et al.** Two epidemiologic patterns of norovirus outbreaks: surveillance in England and Wales, 1992–2000. *Emerging Infectious Diseases* 2002; **9**: 7.