Epidemiol. Infect. (2013), **141**, 1572–1584. © Cambridge University Press 2013
doi:10.1017/S0950268813000496

REVIEW ARTICLE
Heterogeneity in norovirus shedding duration affects community risk

M. O. MILBRATH[1]*, I. H. SPICKNALL[2], J. L. ZELNER[3,4], C. L. MOE[5] AND J. N. S. EISENBERG[2,4]

[1] *Department of Environmental Health Sciences, University of Michigan, Ann Arbor, MI, USA*
[2] *Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA*
[3] *Departments of Sociology and Public Policy, University of Michigan, Ann Arbor, USA*
[4] *Center for the study of Complex Systems, University of Michigan, Ann Arbor, USA*
[5] *Center for Global Safe Water, Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA*

Received 16 August 2012; Final revision 22 January 2013; Accepted 8 February 2013; first published online 18 March 2013

SUMMARY

Norovirus is a common cause of gastroenteritis in all ages. Typical infections cause viral shedding periods of days to weeks, but some individuals can shed for months or years. Most norovirus risk models do not include these long-shedding individuals, and may therefore underestimate risk. We reviewed the literature for norovirus-shedding duration data and stratified these data into two distributions: regular shedding (mean 14–16 days) and long shedding (mean 105–136 days). These distributions were used to inform a norovirus transmission model that predicts the impact of long shedders. Our transmission model predicts that this subpopulation increases the outbreak potential (measured by the reproductive number) by 50–80%, the probability of an outbreak by 33%, the severity of transmission (measured by the attack rate) by 20%, and transmission duration by 100%. Characterizing and understanding shedding duration heterogeneity can provide insights into community transmission that can be useful in mitigating norovirus risk.

Key words: Epidemiology, Norwalk agent and related viruses, risk assessment, Susceptible-Infected-Removed (SIR) model, transmission.

INTRODUCTION

Norovirus is notorious for causing highly explosive epidemics, while also creating significant disease through community endemic transmission. As the most common cause of epidemic gastroenteritis across all age groups [1], noroviruses are responsible for more than 90% of viral gastroenteritis and about 50% of all-cause gastroenteral outbreaks worldwide [2]. In the USA, norovirus infections cause an estimated 71000 hospitalizations annually, costing nearly $500 million per year [3]. The role of endemic norovirus, while not as well characterized as its role in outbreaks, is major; endemic incidence is estimated at around 5% per year for all ages and 20% per year in children aged <5 years [4, 5].

Norovirus is transmitted through aerosolized vomitus and faecal contamination, either directly from person-to-person or indirectly through the

* Author for corresponding: Dr M. O. Milbrath, Environmental Health Sciences, 1415 Washington Heights, Ann Arbor, MI 48104, USA.
(Email: meghanom@umich.edu)

environment. Although there is little data to inform the relative roles of the different transmission pathways, there is evidence to support both symptomatic and asymptomatic transmission [6, 7]. Symptomatic shedding such as a vomiting event has been shown to be the source of many outbreaks, but it is likely that symptomatic individuals remove themselves from the community, limiting their rate of contact. Asymptomatic individuals, on the other hand, mix readily in the public. Additionally, post-symptomatic shedding has been shown to occur in high concentrations [8–10], at levels similar to symptomatic individuals [10].

While typical norovirus infections are self-limiting, with the shedding period lasting from a few days to a few weeks, many individuals can shed noroviruses long past symptom resolution [11–13], and some can shed for months or even years. Infants are often in this long-shedding group; an immature immune system or the presence of maternal antibodies (found in about 75% of infants aged <6 months [14]) can lead to extremely lengthy shedding periods. Similarly, immunocompromised individuals, who are unable to fully clear the infection, may shed norovirus for months or even years [15, 16].

Despite these data indicating the presence of long shedders and the resultant heterogeneity in shedding duration, current norovirus transmission models assume short and homogeneously distributed durations of viral shedding. For example, of two recent models of nosocomial norovirus transmission, one assumes an infectious shedding period of <2 days [13] and the other a distribution from 1 to 8 days [17], while a recent household transmission model estimates an average infectious period of 1·17 days [18]. These short infectiousness periods may be appropriate in understanding transmission in small, enclosed environments such as a house or hospital, where symptomatic individuals shed a high magnitude of virus and quickly exhausts a closely connected and limited population; however, they cannot capture the effects of long shedders in sustaining community transmission.

To better characterize the effect of long-term shedding on norovirus risk, we reviewed the existing literature for empirical data on human norovirus shedding duration. These data are used to estimate distributions of shedding duration that include realistic heterogeneity. This information was then used to inform a norovirus transmission model. Using this model, we demonstrate how the presence of a long-term shedding group affects risk outcomes including the number of infections, duration of transmission, and probability of outbreaks.

METHODS

Literature review search strategy

We reviewed the literature for individual-level human viral shedding duration data through the electronic database Scopus by a key word search using the terms 'norovirus' paired with 'shedding' and/or 'excretion'. These articles were supplemented with sources identified from bibliographies of the resultant studies and with unpublished data from known norovirus researchers.

We restricted the results to English-language human studies containing individual-level data that were acquired through the use of polymerase chain reaction (PCR). Studies that compare PCR with earlier techniques (electron microscopy and ELISA) demonstrate substantial improvements in detection capabilities [8, 19], and thus more accurate estimates of viral shedding duration. PCR does not distinguish between viable and non-viable viruses, and because noroviruses are non-culturable, there is no way to determine the infectivity of those detected through this method [20]. Even with this constraint, PCR data can be used to estimate the duration of infectiousness. It has been shown that passage through a human host does not diminish infectivity [21], meaning that shed viruses, although not culturable, remain viable. Second, norovirus has an extremely small minimum infectious dose [22] – challenge studies have demonstrated infection with 4·8 RT–PCR units [8], and the estimated average probability of infection for a single norovirus particle is about 0·5 [21]. This low infectious dose, along with the high number of particles that are excreted, support the hypothesis that detection indicates a sufficient dose, even if some viral particles have mutated to have non-infectious capsids while remaining sufficiently intact to protect the easily degradable RNA.

Estimating the distribution of shedding durations

In the challenge studies, shedding duration was calculated from the first positive stool sample to the last positive stool sample. For other types of studies (e.g. outbreak investigations), the date of the first positive stool sample was not known, so we assumed that

Table 1. *Criteria used for data stratification*

	Regular shedders	Long shedders
Operational	>1-year-old and competent	Compromised and/or < 1-year-old
Functional	⩽34 days	>34 days

viral shedding starts with the onset of symptoms and continues until the last positive sample. If shedding lengths were reported in days after inoculation [8], we subtracted the number of days to symptom onset to account for the incubation period, and for consistency among the datasets. Although pre-symptomatic shedding has been reported in <30% of the population [23], it is generally for <1 day [8, 24, 25] and there is no indication that it would affect the individual shedding groups differently.

We used two methods to identify the long-shedder population (Table 1). First, we examined the literature to identify individual characteristics potentially related to long-term shedding. If an individual was identified in the original study as having one or more of these characteristics s/he was categorized into the long-shedding group (*operational long shedders*). Because not all individuals in these categories were long shedders, and because others who do not have these characteristics may shed for long times, we developed a comparative method. With the second definition, long shedders were those whose shedding durations exceeded a set maximum shedding length, regardless of individual characteristics (*functional long shedders*). We chose the value of 34 days as this cut-off, as it was the maximum duration in controlled experiments using only healthy individuals [C. L. Moe, previously unpublished data (available in the Supplementary material)] [24, 26]. Because of the uncertainty in these estimates, we also performed a sensitivity analysis on this parameter.

We fitted the data to multiple distributions, and using a likelihood ratio test we found that both the lognormal and gamma distributions fit the data well. We chose gamma distributions based on biological plausibility [27], precedence [28], and because this distribution can be included in a deterministic model using multi-compartmental waiting times when the shape parameter is >1 (see [18] for an explanation). The stratified data were fitted to gamma distributions using the DGAMMA function in R [29], which estimates parameters using maximum likelihood.

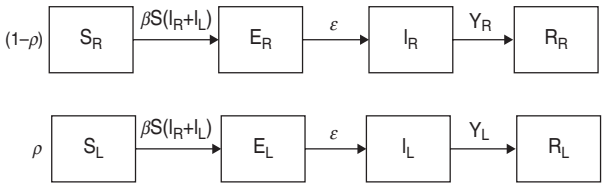

Fig. 1. Model schematic. S, Susceptible; E, exposed; I, infected; R, recovered; subscript ($_R$), regular shedding; subscript ($_L$), long shedding. Parameters: β, transmission probability; ε, 1/incubation period; γ_R, 1/regular-shedder infectious period; γ_L, 1/long-shedder infectious period; ρ, fraction that is long-shedding.

Model design

The data from the review were used to parameterize two modified susceptible (S), exposed (E), infectious (I), recovered (R) transmission models (SEIR models) based on the schematic in Figure 1 and the equations below.

$$\frac{dS_R}{dt} = -\beta S_R(I_L + I_R),$$
$$\frac{dS_L}{dt} = -\beta S_L(I_L + I_R),$$
$$\frac{dE_R}{dt} = \beta S_R(I_L + I_R) - \varepsilon E_R,$$
$$\frac{dE_L}{dt} = \beta S_L(I_L + I_R) - \varepsilon E_L,$$
$$\frac{dI_R}{dt} = \varepsilon E_R - \gamma_R I_R,$$
$$\frac{dI_L}{dt} = \varepsilon E_L - \gamma_L I_L,$$
$$\frac{dR_R}{dt} = I_R \gamma_R,$$
$$\frac{dR_L}{dt} = I_L \gamma_L,$$

We first examined a standard deterministic transmission model that includes both a regular-shedding group (represented by a subscript R) a long-shedding group (represented by a subscript L). Second, we examined a stochastic, discrete-time version of the deterministic model. In both of these models, individuals may be in any of the following states, reflecting the within-host stages of norovirus infection: susceptible (S_R or S_L), exposed/incubating (E_R or E_L), infectious/shedding (I_R or I_L), or recovered (R_R or R_L).

For the deterministic model, vital dynamics such as host births and deaths are excluded because of the relatively short time scale of the simulations (2500 days). We also assumed permanent immunity after recovery. We assumed the absence of competing strains and co-infection, and that host mixing between the regular- and long-shedding groups is proportional.

Table 2. *Parameter definitions and values used in simulations*

	Parameter definition (units)	Mean value* rate (1/days)	Gamma parameters	Source
β	Transmission rate	0·075		
ε	1/length of exposed period	0·76		[13]
γ_{RO}	1/ regular infectious period (operational)	0·061 (1/16·4)	2·2–7·4	Review
γ_{LO}	1/ long-shedding infectious period (operational)	0·009 (1/105·6)	0·8–129·1	Review
γ_{RF}	1/ regular infectious period (functional)	0·069 (1/14·5)	2·7–5·3	Review
γ_{LF}	1/ long-shedding infectious period (functional)	0·007 (1/136·0)	0·7–199·4	Review
ρ	Fraction that is long shedding	0·05		[31, 32]

* Value used when parameter was not varied in simulation.

Infectious states (I_R and I_L) are assumed to exert an equal force of infection due to (1) low infectious dose [21], (2) high levels of asymptomatic excretion [7, 30], and (3) lack of correlation between virus titre in the first positive stool with shedding length [9]. We use a singular infectious period that does not account for symptoms. We examined the more complete process that includes symptomatic transmission and found that our conclusions were robust; even when 50% of infections resulted in a 1-day symptomatic period that was 1000 times more contagious, the importance of long shedding remained (see Supplementary material).

The population is split into two groups: a fixed proportion (ρ) is in the long-shedding group, with the remainder ($1-\rho$) in the regular-shedding group. When time (t) = 0, an infectious seed of one individual is in the regular exposed state (E_R), and the remainder are in the susceptible states (S_R, S_L). Susceptible individuals become exposed (E_R or E_L) at an average per-infected rate of β, and transition into the corresponding infectious/shedding state (I_R or I_L) at an average rate of ε, which is assumed to be the same for both shedding groups. Individuals enter recovered states (R_R and R_L) at rate γ_R for regular shedders, and γ_L for long shedders. Estimated parameters used in these simulations are provided in Table 2.

We estimate the percentage of the U.S. population in the operational long-shedding group, ρ (infants and/or immunocompromised), to be about 5%. The number of infants born in the U.S. in 2007 was 4317119 [31] in a population of 304059724 [32], about 1·4% of the population. The number of immunocompromised individuals in the USA has been previously estimated to be about 10 million individuals (3·6% of the U.S. population), a number that includes organ transplant recipients, HIV-infected individuals, and cancer patients [33].

We assessed risk in the deterministic model by examining changes in the basic reproduction number (R_0), which is a measure of the potential of a disease to spread in a population. R_0 is typically defined as the expected number of secondary infections produced by a single index case in a completely susceptible population [34]. It can be expressed as the product of the expected duration of the infectious period and the rate at which secondary infections occur; in a standard SEIR model with a mean infectious period of $1/\gamma$ and a transmission rate of β, R_0 is given by equation (1).

$$R_0 = \frac{\beta}{\gamma}. \tag{1}$$

In our model, where multiple infectious phases are possible, R_0 is the sum of the expected number of secondary cases generated by an individual in each state, weighted by the probability that the index case will occupy that state. For the model presented here, R_0 is a product of the transmission probability β and the expected time an individual will spend in each of the infectious states I_R and I_L ($1/\gamma_R$ and $1/\gamma_L$) and the probability ρ of being in the long-shedding group [equation (2)].

$$R_0 = \beta\left(\frac{(1-\rho)}{\gamma_R} + \frac{\rho}{\gamma_L}\right). \tag{2}$$

We use a discrete-time stochastic model to examine probability, duration, and severity of outbreaks, where individuals can be in the same disease states as presented in the deterministic formulation above (S_R, S_L, E_R, E_L, I_R, I_L, R_R, R_L; Fig. 1). In this model, we assumed a single population with a frequency-dependent contact rate. The model is initialized with a population of size 10000, with one infected individual at t = 0 and the rest of the population in the susceptible states. The total number of

Table 3. *Description of studies and data used in empirical review*

Study	*N*	Age (years)	Immune status	Genotype	Range (days)	Mean (days)	Median (days)
Atmar & Estes (2008) [8][a]	16	18–50	+	GI.I	8–54	28·6	27
Florescu *et al.* (2008) [36][b]	1	1·8	–		240	240	240
Gallimore *et al.* (2004) [35][c]	1	Child	–	rGII.3	156	156	156
Goller *et al.* (2004) [51][d]	1	18	+	n.r.	8	8	8
Goller *et al.* (2004) [25][e]	9	79–94	+	GII	2–20	10·7	10
Kirkwood & Streitberg (2008) [43][f]	9	0–2	+ (6), – (3)	GII.4, GII.6	2–100	30·8	15
Lee *et al.* (2008) [37][b]	1	0·8	–	GII.2	114	114	114
Leon *et al.* (2011) [24][a]	13	18–48	+	GI.1	8–35	21·3	21
Ludwig *et al.* (2008) [41][g]	9	0–16	–	GII.4, GII.3	22–433	95·7	46
Murata *et al.* (2007) [14][h]	23	0–3	+	n.r.	5–47	17·3	16
Nilsson *et al.* (2003) [38][i]	1	65	–	GII	670	670	670
Parashar *et al.* (1998) [52][j]	1	Adult	+	GII	10	10	10
Schorn *et al.* (2010) [39][k]	9	23–59	–	GII.4, GII.7	97–898	305	230
Siebenga *et al.* (2008) [49][l]	8	0–69	–	GII.4, GIIb–GII.3	35–182	89·3	77
Seitz *et al.* (2011) [26][a]	9	18–50	+	GI.1	4–34	17·6	15
Simon *et al.* (2006) [42][m]	12	0–24	–	n.r.	3–140	40·4	23
Tu *et al.* (2008) [9][e]	14	63–93	+	GII	8–44	25·2	25·5
Westhoff *et al.* (2009) [40][k]	2	73	–	GII	90–165	127·5	127·5
C. L. Moe, unpublished (1)[a]	22	18–50	+	GI.1	8–20	8·3	7
C. L. Moe, unpublished (2)[a]	9	18–50	+	GII.2	3–20	6·9	5

+, Immunocompetent; –, immunocompromised; n.r., not reported; r, recombinant.
[a] Challenge study. [b] Case study of infant transplant recipient. [c] Case study of child with genetic disorder.
[d] Investigation of healthcare worker. [e] Outbreak investigation in an elderly care facility. [f] Prospective hospital study. [g] Retrospective study of paediatric oncology patients. [h] Outpatients at a paediatric clinic. [i] Case study. [j] Outbreak investigation of food handlers. [k] Study of transplant recipients. [l] 2-year hospital survey. [m] Outbreak investigation in paediatric oncology and haematology unit.

new infections Xt is drawn from a binomial distribution (St, λt), where St is composed of both S_R and S_L individuals. The exposed period is assumed to follow an exponential distribution with mean duration $1/\varepsilon$. The regular-shedding and long-shedding infectious periods ($1/\gamma_R$ and $1/\gamma_L$) are assumed to follow gamma distributions with parameters obtained by maximum-likelihood estimation using data collected from our literature review (Table 2).

RESULTS

Data extraction

The literature search resulted in a total of 18 English-language studies with individual-level human norovirus molecular-shedding duration data (listed in Table 3), and we received two previously unpublished datasets (for unpublished data see Supplementary material).

Shedding duration ranged from 2 to 898 days across all studies ($N = 168$ individuals) with total mean and median values of 49 days and 19 days, respectively. Much of the variability in this dataset is explained by stratification of the population on two individual characteristics: compromised immunity and infancy. Sixty-two of the 168 individuals were labelled as immunocompromised and/or infants, thus considered together as *operational long shedders*, while the remaining individuals were considered *operational regular shedders*. Mean shedding lengths between these groups differed by about 90 days (Table 4). The immunocompromised group included a child with a genetic disorder [35], transplant recipients [36–40], paediatric oncology patients [41, 42], and one individual with multiple sustained norovirus infections whose immune status was not evaluated, but was assumed to be immunocompromised [43]. Infants were those individuals whose age was reported as <1 year in the study. It is reasonable to expect differences in shedding length to vary by genotype as well, because of different levels of immune response, but we did not have sufficient data to examine this possibility.

Infants that were immunocompetent shed on average for 22·1 days, 1·3 times longer than those in the

Table 4. *Shedding duration summary from empirical review for operational and functional shedding categories*

	Operational regular shedders	Operational long shedders	Functional regular shedders	Functional long shedders	Total
N (individuals)	106	62	120	48	168
Range (days)	2–54	2–898	2–34	35–898	2–898
Median (days)	13	47·5	13	80	19
Data mean (days)	16·4	105·6	14·5	136·3	49·3
Standard deviation	11·7	155·3	8·5	165·0	103·7
Coefficient of variation	71·3	147·1	58·4	121·1	210·3

regular-shedding group (immunocompetent and >1 year old) who shed for a mean of 16·4 days. Regardless of age, immunocompromised individuals shed on average >6 times longer than regular shedders (139·1 days for >1 year old and 106·6 days for infants). The observed effect of both compromised immunity and infancy (90·2 days) is less than the expected additive joint effect (128·4 days). We tested for additivity using linear regression in R [29]; the *P* value for this test was <0·01, prompting us to reject the additive effects model. This indicates that either condition can cause long shedding, but both together (being both an infant and immunocompromised) does not result in extremely long-shedding durations beyond the effect of either condition alone.

This operational definition is useful for the *a priori* identification of high-risk individuals and for understanding the causes underlying long shedding, but was not ideal for our analysis as many cases in our dataset are missing data on age or immune status, and thus cannot be correctly classified using our operational criteria. Further, some individuals who are labelled as immunocompromised, and therefore labelled as long shedders with our operational definition, actually shed for short durations. We therefore, stratified the data based on actual shedding length (*functional long shedders* shed for >34 days). The mean of the two groups by the functional definition differs by over 120 days (Table 4).

The fitted distribution for the unstratified data is highly variable (Fig. 2). When we stratify the data based on both of our defined shedding groups (functional and operational), much of this variability remains in the long-shedding group (coefficient of variation: 147·1 operational, 121·1 functional), while the regular-shedding groups are more homogenous (coefficient of variation: 71·3 operational, 58·4 functional) (Table 4). Using the law of total variance, the variance of the aggregated populations is 10826·2 for the operational definition and 20989·2 for the functional definition. By dividing the variance of the means by the total variance, we estimated that stratification into long-shedding and regular-shedding groups explains 35% (functional) to 37% (operational) of the total variance.

Deterministic model results

Changes in R_0 determine whether an epidemic will occur as well as the subsequent course of transmission. If $R_0 \leqslant 1$, transmission will be limited to sporadic secondary infections, and a major outbreak will not occur, whereas if $R_0 > 1$ an outbreak can result. Based on equations (1) and (2), and assuming $\rho = 5\%$, the inclusion of a long-shedding group [mean = 105·6 days (operational) and 136·3 days (functional)], increases risk by increasing R_0 (1·3-fold increase for the operational definition, 1·5 for functional). Because our value of ρ is estimated, we examined the effect of its increase; if ρ grows to 10% of the population (e.g. increased survival rates for those with compromised systems or an increase in birth rate), the inclusion of the long-shedding group results in a 1·5-fold (operational) or 1·7-fold (functional) increase in R_0 over a population without long shedders. Similarly, a twofold increase in the mean shedding length of the long-shedding group ($1/\gamma_L$) results in a 1·6-fold (operational) or 1·8-fold (functional) increase in the value of R_0 (Table 5). The effects of changes in these long-shedding related values are independent of the transmission rate, β, meaning that resultant relative increases in R_0 are robust to scenarios of different transmission rates, including environmental or symptomatic transmission.

Stochastic model results

Results from stochastic simulations demonstrate that the presence of long shedders increases the probability that an outbreak will occur while increasing the

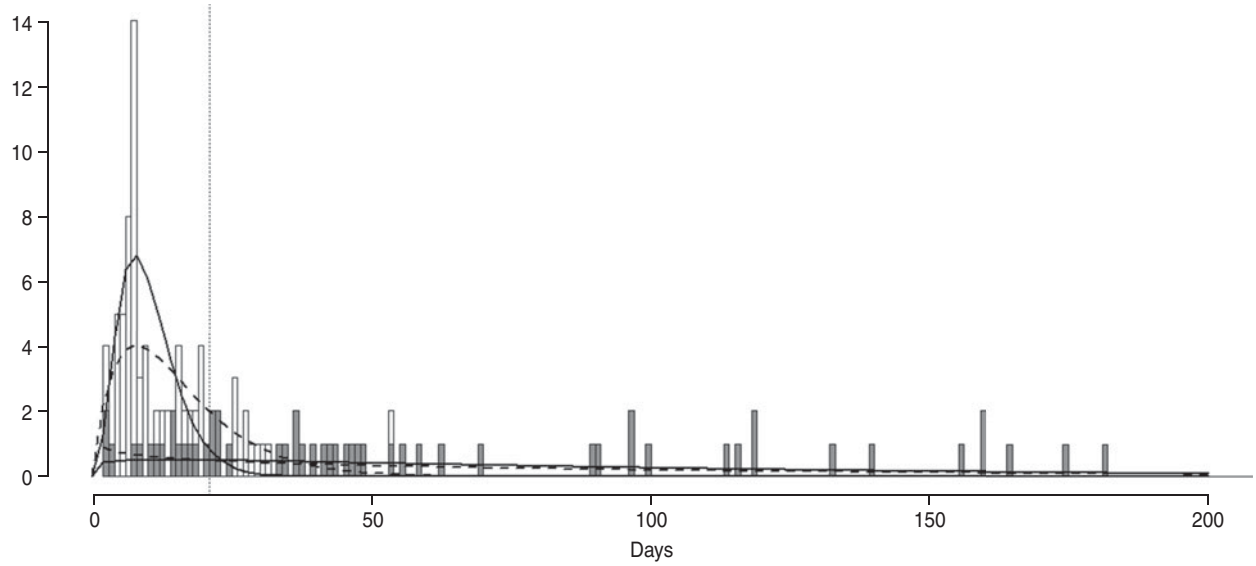

Fig. 2. Shedding lengths and fitted gamma distributions of individual shedding duration data from empirical review. Data are truncated at 200 days; the figure does not include seven individuals who shed for >200 days (up to 898 days). Grey bars indicate operational long shedders (infants and immunocompromised individuals) and white bars indicate operational regular shedders (immunocompetent non-infants). The dotted verticalline at 34 days delineates the cut-off point for the functional definition. Dashed lines represent fitted gamma distributions for the operational definition (regular-shedder parameters = 2·2, 7·4; long-shedder parameters = 0·8, 129·1) and solid lines are gamma distributions fitted to the functionally defined populations, where regular shedders <34 days, long-shedders ⩾34 days (regular-shedder parameters = 2·7, 5·3; long-shedder parameters = 0·7, 199·4).

Table 5. *Deterministic model results: estimated fold increase in R_0 with addition of long shedders*

	$\rho = 0{\cdot}05$, estimated γ_L	$\rho = 0{\cdot}05$, estimated γ_L*2	$\rho = 0{\cdot}10$, estimated γ_L
Functional definition	1·3	1·6	1·5
Operational definition	1·5	1·8	1·7

severity and duration of transmission when disease does not die out. We examined changes in the probability of high transmission events where high transmission was defined as >200 cases. Based on our simulation results, transmission above this level was sustained long term, and below this level would die out. In the absence of long shedders ($\rho = 0$, i.e. the entire population draws shedding durations from the functional regular-shedding distribution), there is a 23% chance that high transmission will occur in 1000 runs using our estimated parameters (Table 6). If we allow 5% of the population to draw from the functional long-shedding distribution, the probability of high transmission increases to 31%. This effect is less evident using the operational definition, where the probability remains at about 37% (Table 6). As shown in Figure 3, including a long-shedding group by either definition increases both the attack rate, and its duration of sustained transmission. With the addition of long shedders, the mean attack rate increases 145% (operational) to 225% (functional) over a population with only regular shedders (Table 6). The duration of sustained transmission also increases, approximately doubling for both definitions (operational: 633–1199 days; functional: 756–1683days). Figure 3 demonstrates that these outcomes are dependent on our method of stratification. For example, the number of cases in the absence of long shedders is much higher using the operational definition (filled circles) than the functional definition (open circles), due to the wider range and higher mean of this subset.

Sensitivity analysis

Our estimates are based on an incomplete dataset of all true shedding times ($1/\gamma_L$) due to publication bias and the difficulty of obtaining data on asymptomatic community infection. We therefore examined the sensitivity of our outcomes over a wide range of values

Table 6. *Results of stochastic simulations using parameters from the review*

	Probability of outbreak (%)	Mean attack rate (%)	Fold increase in outbreak duration
Operational – no long shedders	36	45·8	—*
Operational – with long shedders	38	67·1	1·9
Functional – no long shedders	23	29·4	1·2
Functional – with long shedders	31	66·3	2·7

* Set as baseline, column indicates fold increase over this value.

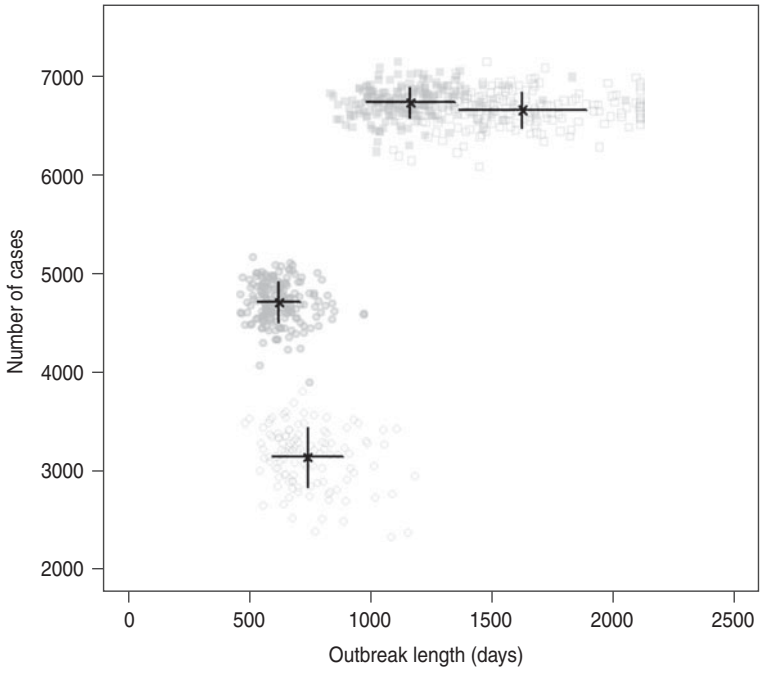

Fig. 3. Outbreak duration and severity using shedding length parameter values from our empirical review (Table 2). Only model runs resulting in high transmission events (>200 cases) are shown. Runs with both long-shedding and regular-shedding groups are represented by squares, runs with only regular shedders are represented by circles. Filled shapes indicate use of the operational (*a priori*) definition and open shapes indicate use of the functional definition (>34 days) for long shedders. Black lines show the mean and standard distribution for each scenario.

for this parameter. The probability of high transmission increases by about 10% as the mean long-shedding duration ($1/\gamma_L$) increases from 20 to 250, levelling off at a maximum probability of 42% for the operational definition and 34% for the functional definition when $\gamma_L > 100$ days (Fig. 4). The higher probability in the operational definition is again attributed to the wider range and higher mean shedding duration of the regular-shedding group, which contains the bulk (95%) of the population.

The total number of infected individuals also increases with $1/\gamma_L$, as shown in Figure 5. In simulations in a population of 10000 individuals, the maximum number of cases is about 8000 for both definitions (80% of the population), and remains well above the average number of cases in the absence of long shedders (46% operational, 29% functional). Transmission duration, measured as the last time at which there is at least one infected individual in the population, increases as $1/\gamma_L$ increases (Fig. 6).

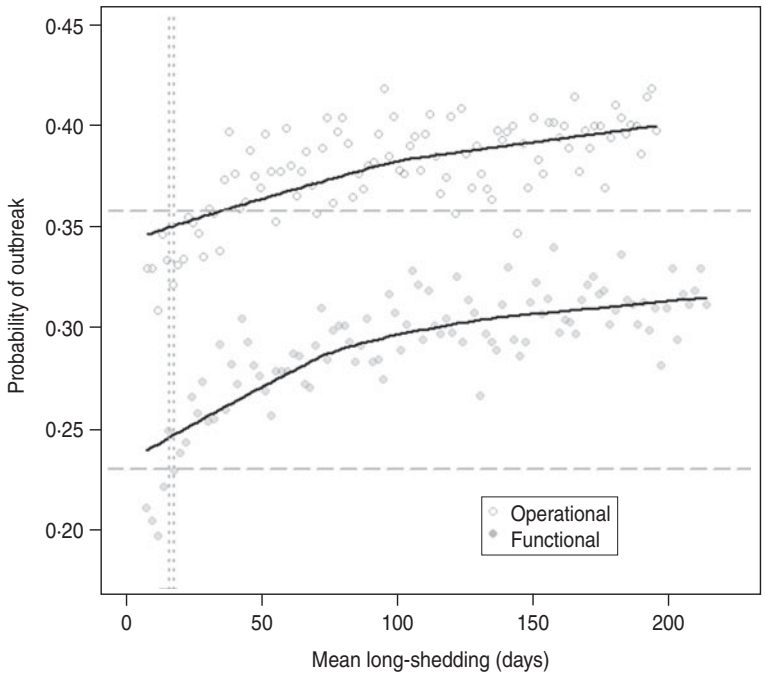

Fig. 4. Sensitivity of outbreak probability to variable long-shedding duration ($1/\gamma_L$). Outbreaks are defined by having >200 cases. Dashed lines show the probability for each definition in the absence of long shedders, and dotted lines represent the shedding duration of the regular-shedding groups ($\gamma = 16{\cdot}4$ days, $\gamma = 14{\cdot}5$ days). Open circles represent use of the operational definition, and filled grey circles represent use of the functional definition for stratification.

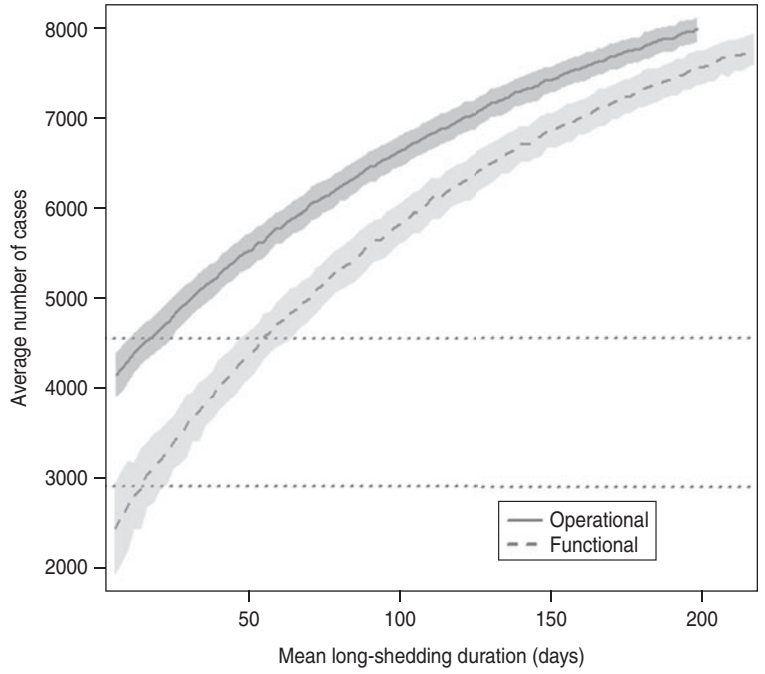

Fig. 5. Sensitivity of outbreak severity to variable long-shedding duration ($1/\gamma_L$). Outbreaks are defined by having >200 cases. Dashed line shows average and standard definition for 1000 runs with the functional definition (solid line) shows average and standard deviation for operational definition. Dotted lines show average number of cases in the absence of long shedders.

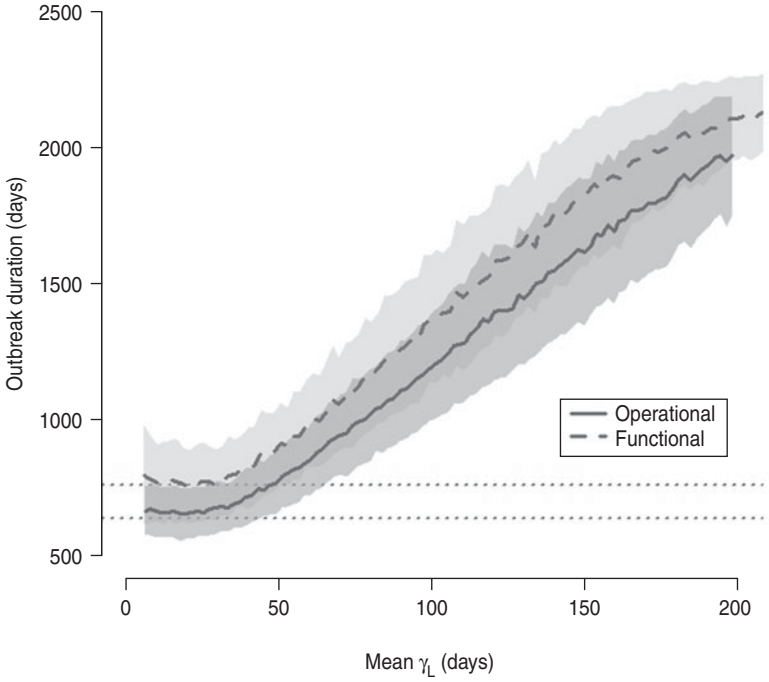

Fig. 6. Sensitivity of outbreak duration to variable long-shedding duration ($1/\gamma_L$). Outbreaks are defined by having >200 cases. End of outbreak occurs when there are no exposed or infected individuals remaining. Dotted lines show average duration for each definition in the absence of long shedders.

The maximum expected duration of about 2000 days (about 5·5 years) is reached when $1/\gamma_L = 200$ days for both definitions. Again, this is much longer than the average duration in the absences of long shedders (633 days operational, 756 days functional).

DISCUSSION

Heterogeneity in viral shedding duration has clear implications for norovirus transmission dynamics, and should be included when estimating disease risks. In the case of norovirus, estimates of short average shedding duration do not capture the potential effects of long-term shedding individuals in sustaining community transmission. Our empirically derived characterization of norovirus shedding duration allows us to model its impact on population-level norovirus risk. Our analysis indicates that inclusion of long shedders results in a 50–80% increase in the basic reproduction number R_0, about a 20% increase in the number of cases, and about a 100% increase in the duration that transmission is sustained in a population.

The importance of an increase in R_0 depends on its original value, as the number of excess cases that result from an increase in R_0 is highly dependent on the transmission rate, β. For example, if R_0 is already quite high (such as within a home), then an outbreak would already be likely to spread through a population and a twofold increase would not have much effect. If $R_0 \leqslant 1$ (probably true for community transmission), then this same increase can mean the difference of an outbreak dying out or of it taking off in a population (see [44, fig. 1] for further explanation). Long shedders are likely, therefore, to be more important for community transmission where they can potentially impact the probability of outbreak occurrence.

Since long shedders tend to be asymptomatic, estimating the transmission rate due to asymptomatic shedders is important. If asymptomatic individuals do not transmit norovirus then long shedders cannot play a role in norovirus transmission. However, evidence and documentation that asymptomatic individuals can be the index case of outbreaks suggest that asymptomatic transmission does occur and is probably enhanced by long shedders [6, 7].

Any effect of long shedders would become more profound with the addition of more complicated and realistic structure, including social structure and effect of varied levels of infectiousness. In a structured metapopulation comprised of loosely connected groups,

a pathogen must persist within its initial group long enough to allow for migration into another one. This type of long-shedder-driven persistence may provide an explanation for the explosive but episodic character of norovirus outbreaks: long shedders may allow the virus to circulate at low levels in the population until it either reaches an individual with a high contact rate (e.g. a food handler or symptomatic individual in a closed scenario), or a pocket of susceptible individuals (e.g. an unexposed school), resulting in a large new outbreak. A socially structured model could also account for other transmission processes, including clustering of long shedders in healthcare facilities and other high-risk zones.

The effect of long shedders may also become more profound if we relax the assumption of homotypic immunity. Immunity has been estimated to last 14 weeks [45], but this is a highly debated topic [46]. When the duration of shedding ($1/\gamma_L$) is greater than the duration of immunity, the long-shedding group can act as a reservoir, allowing the infection to persist in a population and potentially re-seed periodic epidemics, in a manner similar to herpes viruses and other pathogens with long-term carriage, such as *Salmonella typhi*.

Because host immune pressure can impact calicivirus RNA evolution [38], long shedders may also impact the population ecology of noroviruses [47]. Since the amino acids that mutate most frequently are those involved in immune evasion [48, 49], escape mutants may emerge over the life of a single long-shedding infection. In fact, the number of amino-acid mutations arising over the course of a year within-host is similar to the number distinguishing outbreak variants from each other [49]. Consequently, long shedders represent a potential mechanism for introducing novel variants into populations that have achieved herd immunity against the most recent circulating strains.

Our parameter estimates are sensitive to differences in technique and study design across studies (Table 1), including differences in faecal viral concentration, specimen storage, RNA extraction efficiency, presence of faecal reverse transcriptase inhibitors, and primer usage [50]. Many studies terminated sample collection while patients were still shedding [8, 38, 51], resulting in right-censored data, and systematic underestimates of actual shedding durations. Publication bias exists in the literature, so that published studies may not reflect the true distribution in the population. Future studies, with increasingly sensitive techniques, will provide more accurate estimates of actual post-infection viral shedding duration.

CONCLUSIONS

We have demonstrated the importance of including an empirically validated representation of between-host heterogeneity in norovirus natural history when assessing population-level transmission risks. These results should be incorporated in models that include other potentially important sources of transmission heterogeneity such heterogeneous contact networks, protective immunity, and strain competition. This will facilitate the development of policies and interventions that can target the individuals who are both the most susceptible and the most likely to transmit disease. As we deepen our understanding of the different types, degrees, and interactions of heterogeneity in disease transmission systems, we can make more informed policy decisions and recommendations and more effectively protect human health.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/S0950268813000496.

ACKNOWLEDGEMENTS

This work was funded under the U.S. Environmental Protection Agency Science to Achieve Results (STAR) grant no. RD83172701 and by the U.S. Department of Homeland Security University Programs grant no. R83236201.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Fankhauser RL, *et al.*** Epidemiologic and molecular trends of 'Norwalk-like viruses' associated with outbreaks of gastroenteritis in the United States. *Journal of Infectious Diseases* 2002; **186**: 1–7.
2. **Patel MM, *et al.*** Noroviruses: a comprehensive review. *Journal of Clinical Virology* 2009; **44**: 1–8.
3. **Lopman BA, *et al.*** Increasing rates of gastroenteritis hospital discharges in US adults and the contribution of norovirus, 1996–2007. *Clinical Infectious Diseases* 2011; **52**: 466–474.

4. **Lopman B, *et al.*** Environmental transmission of norovirus gastroenteritis. *Current Opinion in Virology* 2012; **2**: 96–102.
5. **Tam CC, *et al.*** Changes in causes of acute gastroenteritis in the United Kingdom over 15 years: microbiologic findings from 2 prospective, population-based studies of infectious intestinal disease. *Clinical Infectious Diseases* 2012; **54**: 1275–1286.
6. **Sukhrie FHA, *et al.*** Chronic shedders as reservoir for nosocomial transmission of norovirus. *Journal of Clinical Microbiology* 2010; **48**: 4303–4305.
7. **Ozawa K, *et al.*** Norovirus infections in symptomatic and asymptomatic food handlers in Japan. *Journal of Clinical Microbiology* 2007; **45**: 3996–4005.
8. **Atmar RL, *et al.*** Norwalk virus shedding after experimental human infection. *Emerging Infectious Diseases* 2008; **14**: 1553–1557.
9. **Tu ETV, *et al.*** Norovirus excretion in an aged-care setting. *Journal of Clinical Microbiology* 2008; **46**: 2119–2121.
10. **Aoki Y, *et al.*** Duration of norovirus excretion and the longitudinal course of viral load in norovirus-infected elderly patients. *The Journal of hospital infection* 2010; **75**: 42–6.
11. **Gallimore CI, *et al.*** Asymptomatic and symptomatic excretion of Noroviruses during a hospital outbreak of gastroenteritis. *Journal of Clinical Microbiology* 2004; **42**: 2271–2274.
12. **Rockx B, *et al.*** Natural history of human Calicivirus infection: a prospective cohort study. *Clinical Infectious Diseases* 2002; **35**: 246–253.
13. **Vanderpas J, *et al.*** Mathematical model for the control of nosocomial norovirus. *Journal of Hospital Infection* 2009; **71**: 214–222.
14. **Murata T, *et al.*** Prolonged norovirus shedding in infants $\leqslant$6 months of age with gastroenteritis. *Pediatric Infectious Disease Journal* 2007; **26**: 46–49.
15. **Henke-Gendo C, *et al.*** New real-time PCR detects prolonged norovirus excretion in highly immunosuppressed patients and children. *Journal of Clinical Microbiology* 2009; **47**: 2855–2862.
16. **Furuya D, *et al.*** Age, viral copy number, and immunosuppressive therapy affect the duration of norovirus RNA excretion in inpatients diagnosed with norovirus infection. *apanese Journal of Infectious Diseases* 2011; **64**: 104–108.
17. **Lee BY, *et al.*** Economic impact of outbreaks of norovirus infection in hospitals. *Infection Control and Hospital Epidemiology* 2011; **32**: 191–193.
18. **Zelner JL, *et al.*** How infections propagate after point--source outbreaks: an analysis of secondary norovirus transmission. *Epidemiology* 2010; **21**: 711–718.
19. **Amar CFL, *et al.*** Detection by PCR of eight groups of enteric pathogens in 4,627 faecal samples: re-examination of the English case-control Infectious Intestinal Disease Study (1993–1996). *European Journal of Clinical Microbiology and Infectious Diseases* 2007; **26**: 311–323.
20. **Li D, *et al.*** Critical studies on binding-based RT-PCR detection of infectious Noroviruses. *Journal of Virological Methods* 2011; **177**: 153–9.
21. **Teunis PFM, *et al.*** Norwalk virus: how infectious is it? *Journal of Medical Virology* 2008; **80**: 1468–1476.
22. **Yezli S, Otter JA.** Minimum infective dose of the major human respiratory and enteric viruses transmitted through food and the environment. *Food and Environmental Virology* 2011; **3**: 1–30.
23. **Glass RI, Parashar UD, Estes MK.** Norovirus gastroenteritis. *New England Journal of Medicine* 2009; **361**: 1776–1785.
24. **Leon JS, *et al.*** Randomized, double-blinded clinical trial for human norovirus inactivation in oysters by high hydrostatic pressure processing. *Applied and Environmental Microbiology* 2011; **77**: 5476–5482.
25. **Goller JL, *et al.*** Long-term features of norovirus gastroenteritis in the elderly. *Journal of Hospital Infection* 2004; **58**: 286–291.
26. **Seitz SR, *et al.*** Norovirus human infectivity and persistence in water. *Applied and Environmental Microbiology* 2011; **77**: 6884–6888.
27. **Höhne M, Schreier E.** Detection and characterization of norovirus outbreaks in Germany: application of a one-tube RT-PCR using a fluorogenic real-time detection system. *Journal of Medical Virology* 2004; **72**: 312–319.
28. **Sartwell PE.** The incubation period and the dynamics of infectious disease. *American Journal of Epidemiology* 1966; **83**: 204–216.
29. **R Development Core Team.** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2010.
30. **Marshall JA, *et al.*** High level excretion of Norwalk-like virus following resolution of clinical illness. *Pathology* 2001; **33**: 50–52.
31. **Hamilton B, Martin J, Ventura S.** Preliminary data for 2007. *National Vital Statistics Reports* 2009; **57**.
32. **US Census Bureau.** State and county quick facts, 2009 (http://quickfacts.census.gov/qfd/states/00000.html). Accessed: 26 October 2010.
33. **Kemper AR, Davis MM, Freed GL.** Expected adverse events in a mass smallpox vaccination campaign. *Effective Clinical Practice* 2002; **5**: 84–90.
34. **Anderson RM, May RM.** *Infectious Diseases of Humans Dynamics and Control*. Oxford: Oxford University Press, 1992.
35. **Gallimore CI, *et al.*** Chronic excretion of a norovirus in a child with cartilage hair hypoplasia (CHH). *Journal of Clinical Virology* 2004; **30**: 196–204.
36. **Florescu DF, *et al.*** Two cases of Norwalk virus enteritis following small bowel transplantation treated with oral human serum immunoglobulin. *Pediatric Transplantation* 2008; **12**: 372–375.
37. **Lee BE, *et al.*** Chronic norovirus and adenovirus infection in a solid organ transplant recipient. *Pediatric Infectious Disease Journal* 2008; **27**: 360–362.
38. **Nilsson M, *et al.*** Evolution of human calicivirus RNA in vivo: accumulation of mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a new phenotype. *Journal of Virology* 2003; **77**: 13 117–13 124.

39. **Schorn R, *et al.*** Chronic norovirus infection after kidney transplantation: molecular evidence for immune-driven viral evolution. *Clinical Infectious Diseases* 2010; **51**: 307–314.
40. **Westhoff TH, *et al.*** Chronic norovirus infection in renal transplant recipients. *Nephrology Dialysis Transplantation* 2009; **24**: 1051–1053.
41. **Ludwig A, *et al.*** Quantitative detection of norovirus excretion in pediatric patients with cancer and prolonged gastroenteritis and shedding of norovirus. *Journal of Medical Virology* 2008; **80**: 1461–1467.
42. **Simon A, *et al.*** Norovirus outbreak in a pediatric oncology unit. *Scandinavian Journal of Gastroenterology* 2006; **41**: 693–699.
43. **Kirkwood CD, Streitberg R.** Calicivirus shedding in children after recovery from diarrhoeal disease. *Journal of Clinical Virology* 2008; **43**: 346–348.
44. **Koopman JS, Longini IM.** The ecological effects of individual exposures and nonlinear disease dynamics in populations. *American Journal of Public Health* 1994; **84**: 836–842.
45. **Matsui SM, Greenberg HB.** Immunity to calicivirus infection. *Journal of Infectious Diseases* 2000; **181** (Suppl. 2): S331–S335.
46. **Donaldson EF, *et al.*** Viral shape-shifting: Norovirus evasion of the human immune system. *Nature Reviews Microbiology* 2010; **8**: 231–241.
47. **Grenfell BT, *et al.*** Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 2004; **303**: 327–332.
48. **Lindesmith LC, *et al.*** Mechanisms of GII.4 norovirus persistence in human populations. *PLoS Medicine* 2008; **5**: 269–290.
49. **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.
50. **Patel MM, *et al.*** Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerging Infectious Diseases* 2008; **14**: 1224–1231.
51. **Goller JL, *et al.*** Norovirus excretion in a healthcare worker without major symptoms of gastroenteritis: Infection control implications. *Australian and New Zealand Journal of Public Health* 2004; **28**: 88–89.
52. **Parashar UD, *et al.*** An outbreak of viral gastroenteritis associated with consumptions of sandwiches: implications for the control of transmission by food handlers. *Epidemiology and Infection* 1998; **121**: 615–621.