Norovirus is the leading cause of outbreaks of acute gastroenteritis (AGE) in the United States (1,2). The Centers for Disease Control and Prevention (CDC) collects data on AGE outbreaks through the National Outbreak Reporting System (NORS). During 2009–2017, norovirus was the suspected or confirmed etiology of 47% of AGE outbreaks reported to NORS (3). The size and severity of outbreaks varies across different settings, times of year, and genotypes, suggesting norovirus transmissibility is variable across different outbreak settings and contexts (4). Generally, the transmission potential of infectious diseases is influenced by the infectiousness of the pathogen, the duration of infectiousness, and the number of susceptible contacts exposed during the infectious period (5).

The reproduction number is a metric for quantifying transmissibility of a pathogen. The basic reproduction number ($R_0$) is the average number of secondary cases that arise from a primary case in a completely susceptible population. The effective reproduction number ($R_e$) quantifies the average number of secondary cases that arise from a primary case in a population that is not completely susceptible. $R_e$ varies over the course of an outbreak as the proportion of the susceptible population changes (6,7). $R_0$ and $R_e$ are not just metrics of the biologic properties of pathogens but also measures of the transmissibility of a pathogen within a specific population or setting (8,9).

Several transmission modeling studies in different settings have estimated $R_0$ and $R_e$ of norovirus, but a large variation in these estimates occurs and $R_0$ ranges from 1.1–7.2 (10). Much of the $R_0$ variation likely is due to differences in the structures, population mixing assumptions, and data between transmission models in different settings (10). Generally, model estimates from community surveillance data result in an $R_0$ of $\approx$2, but estimates from outbreak data tend to be higher and more variable. The variability of estimates from models that use outbreak data likely are driven by context; outbreaks might occur in populations that are not representative of the population as a whole and transmission likely is higher in these settings than in the community (4).

We estimated $R_0$ and $R_e$ for thousands of norovirus outbreaks in the United States. We evaluated whether $R_0$ was associated with setting, season, year, or geographic region. In addition, we assessed whether norovirus was suspected or confirmed as the cause of the outbreak.

**Methods**

**Data**

We obtained data on all norovirus outbreaks during 2009–2017 from NORS and CaliciNet (11). We defined...
an outbreak as ≥2 epidemiologically linked cases of suspected or laboratory-confirmed norovirus. NORS data consist of web-based reports of all foodborne, waterborne, and enteric disease outbreaks transmitted by contact with environmental sources, infected persons or animals, or unknown modes of transmission reported by state, local, and territorial public health agencies. This web-based reporting system collects epidemiologic information, including the dates; settings, such as long-term care facilities, child daycare facilities, hospitals, and schools; geographic location of the outbreak; the estimated total number of cases; and exposed population (2). For settings that report staff and guest case numbers, we included these data in the estimated total number of cases and exposed population. CaliciNet data consists of sequence-derived genotypes and epidemiologic data from norovirus outbreaks submitted from local, state, and federal public health laboratories. We obtained CaliciNet genotypes that were linked to outbreak data we acquired from NORS.

For all outbreaks reported to NORS, data are collected on the total estimated primary cases, including all laboratory-confirmed and suspected primary cases. These data exclude cases associated with secondary illnesses, such as person-to-person norovirus transmission in households after a restaurant-based outbreak. However, data for calculating attack rates, specifically the number of exposed persons and the subset of the exposed persons who became ill, are only collected for outbreaks with person-to-person, environmental, or unknown transmission modes. In addition, data collected from outbreaks might not be documented consistently across a report. For example, outbreaks for which setting-specific information on the total number of guests and staff that are reported to be ill, referred to as total ill, might not match the reported total estimated primary cases. During 2009–2017, a total of 17,822 suspected and confirmed norovirus outbreaks were reported to NORS. We excluded 10,728 outbreaks based on the following criteria, which we imposed hierarchically: transmission was not person-to-person (n = 3,866); the outbreak exposure occurred in multiple states (n = 8); the outbreak occurred in Puerto Rico (n = 3), which we excluded because of small sample size; the size of total exposed population or major setting were not reported (n = 5,573); the total estimated primary cases and the total ill among the exposed population were not equal (n = 1,231); or the total estimated primary cases or the total ill among the exposed population were reported to be greater than the total exposed population size (n = 47) (Appendix Figure 1, https://wwwnc.cdc.gov/EID/article/26/8/19-1537-App1.pdf). In all, 7,094 norovirus outbreaks met our inclusion criteria in subsequent analyses (Appendix Table 1). We did not use imputation techniques to infer values for missing data because no good proxy variables inferred missing data for major settings and exposed population size.

**Estimating R₀ and Rₚ**

We used the final size method to calculate R₀, Rₚ, and associated SEs (12; Appendix). The final size method calculates R₀ and Rₚ based on 3 variables: the total population size of the outbreak (N), the total number of cases in the outbreak (C), and the number of susceptible persons at the start of the outbreak (S). In our calculations, C was informed by NORS outbreak data for the estimated total number ill and N by the exposed population. NORS data does not include nor can it inform the number of susceptible persons at the start of an outbreak. Therefore, to estimate S, we used norovirus challenge study data on the percent of persons that become infected and develop AGE after challenge with virus. Across all published studies, the weighted average of participants in whom gastroenteritis developed after challenge is 47% (range 27%–80%; Appendix Table 1) (13–19). We assumed S is the number of persons susceptible to disease, as opposed to infection. To calculate S, we multiplied 47% by N and rounded to the nearest integer. For 890 outbreaks, the total number of cases, C, was greater than our estimated S; for these outbreaks we set S equal to C, corresponding to a 100% attack rate. We also calculated S assuming 27% and 80% of N were susceptible to assess the sensitivity of our model results to this parameter.

**Regression Analysis**

After estimating R₀, Rₚ, and associated SEs for each norovirus outbreak, we fit a linear regression model to the log-transformed estimated reproduction numbers to assess whether outbreak setting, census region, season, year, suspected or confirmed norovirus, or genotype were associated with transmissibility. All variables were categorical, where the reference was assigned as the group with the most outbreaks reported, except for the suspected or confirmed variable, for which we set the referent to outbreaks with confirmed norovirus etiology. We used weighted least squares combined with estimated standard errors to produce robust estimates accounting for heteroscedasticity and non-normally distributed model residuals by using the estimatr package in R version 3.4.2 (20,21). We included the following variables in our models: outbreak setting; census region;
meteorological season, defined as spring (March 1–May 31), summer (June 1–August 31), fall (September 1–November 30), or winter (December 1–February 28); year, defined as July–June; whether norovirus was suspected or confirmed; and norovirus genotype, categorized as GI, GII.4, or GII.non4.

For outbreaks for which we calculated $R_0$ and $R_e$, we had norovirus genotype data for only 22% (1,571). In a preliminary analysis, we fit a univariate linear regression model to estimate $R_0$ by norovirus genotype alone and by norovirus genotype and year and found no evidence for variation (Appendix). Given these results and the small sample size, we did not include norovirus genotype in our models and performed model selection on the remaining variables. To determine which variables to include, we used a forward selection process and selected the model with the lowest Akaike information criterion and Bayes information criterion values.

**Sensitivity Analysis**

We tested the sensitivity of our regression model results to different modeling approaches and different assumptions of the percent susceptible at the start of an outbreak. We also fit a logistic regression model of binary transmission and a negative binomial regression of the final outbreak size by using the log-transformed exposed population size as a measure of the attack rate of an outbreak. Thus, we could make comparisons between the models to see if the results from modeling continuous transmission were consistent with the results of modeling binary transmission and attack rates. In addition, we ran all the regression models again using the assumption that 27% and 80% susceptible at the start of an outbreak, which corresponds to the minimum and maximum percent susceptible to AGE from published challenge studies (Appendix Table 2).

**Results**

Of the 7,094 norovirus outbreaks included in our final dataset, 75% (5,335) occurred in long-term care and assisted living facilities and 57% (4,016) occurred in winter. The median outbreak size was 28 cases (interquartile range [IQR] 16–47) and the median attack rate was 22% (IQR 11%–36%) (Table 1). The median $R_0$ was 2.75 (IQR 2.38–3.65) and the median $R_e$ was 1.29 (IQR 1.12–1.74).

**Model Selection and Regression Analysis**

The final selected model included the following variables: major setting, census region, season, year, and whether norovirus was suspected or confirmed (Akaike information criterion = 5,803; Bayes information criterion = 5,968) (Appendix Table 3). For long-term care and assisted living facilities, $R_0$ was 3.35 (95% CI 3.26–3.45). $R_e$ for outbreaks in all other settings did not differ substantially, except for outbreaks in schools, colleges, and universities, in which $R_0$ was slightly reduced, 2.92 (95% CI 2.82–3.03) (Table 2; Appendix Figure 2). We found that $R_e$ differed substantially by outbreak status; suspected norovirus outbreaks had a lower $R_{e\rho}$ 3.02 (95% CI 2.94–3.10), than that for confirmed outbreaks ($R_e = 3.35$ [95% CI 3.26–3.45]).

Estimated $R_0$ varied only slightly by census region and was lowest in the northeast ($R_0 = 3.00$ [95% CI 2.92–3.08]). Season and year also contributed to changes in the $R_0$. Estimated $R_0$ was highest in winter (3.35 [95% CI 3.26–3.45]) and fall (3.37 [95% CI 3.24–3.50]) and lowest during the summer months (3.11 [95% CI 2.97–3.25]). Outbreaks reported during January 2009–June 2012 all had higher estimated $R_0$ (range for individual seasonal years 3.77–3.93) than the reference period, July 2016–June 2017 (Table 2; Appendix Figure 2). Our findings were generally robust to assumptions about the proportion susceptible at the start of the outbreak and whether we modeled the outcome of $R_0$, $R_{e\rho}$, or final outbreak size (Appendix Tables 4–6, Figure 3).

**Discussion**

By using a large national outbreak dataset, we investigated transmission patterns of norovirus outbreaks. Our analysis led to several key findings. First, reported norovirus outbreaks in the United States have modest $R_0$ (2.75 [IQR 2.38–3.65]) and $R_e$ (1.29 [IQR 1.12–1.74]) values. Second, we found that $R_0$ and $R_e$ did not vary across most settings, except for outbreaks in schools, colleges, and universities, which had lower estimated transmission values. Third, we found higher transmission in laboratory-confirmed outbreaks relative to suspected outbreaks and higher transmission for outbreaks occurring in the winter months relative to summer months.

Our finding that norovirus outbreaks in the United States have modest transmission values is somewhat surprising. In a recent review of norovirus modeling studies, Gaythorpe et al. (10) found $R_0$ estimates for norovirus were 1.1–7.2. Of note, $R_0$ and $R_e$ estimates from transmission modeling studies that analyzed data from norovirus outbreaks were high, but variability between studies was high; $R_0$ estimates were ≈1–14 (22–24). Our estimates are within the reproduction numbers estimated by using transmission models of norovirus based on outbreak data (22,25).
Characterizing Norovirus Transmission

However, our estimates are higher than those from several studies that estimated reproduction numbers by using population-level transmission models (26–29), suggesting that transmission of norovirus in outbreak settings is higher than sporadic transmission in the community.

From our main analysis, we found that outbreaks in schools, colleges, and universities had lower estimated transmission, but transmission varied little across all other settings. Relative to outbreaks in long-term care and assisted living facilities, outbreaks that occurred in private homes or residences and restaurants had higher final sizes, and schools, colleges, and universities had lower estimated attack rates. Our finding that outbreaks in the winter had higher estimated transmissibility than outbreaks that occurred in summer is likely a factor of the strong wintertime seasonality of noroviruses in the United States (30,31). Consistent with this finding are the observations that norovirus case and outbreak reports are inversely correlated with temperature (30,31) and that survival of norovirus surrogate viruses, such as murine norovirus and feline calicivirus, declines with increasing temperatures (32,33).

Several differences we found might be driven by surveillance biases rather than differences in norovirus transmission. Suspected norovirus outbreaks without a laboratory-confirmed outbreak etiology had lower transmission than laboratory-confirmed norovirus outbreaks, perhaps because suspected outbreaks are not investigated as well as confirmed outbreaks and have lower rates of case ascertainment. Outbreaks reported in the south had higher estimated $R_0$ and $R_e$ relative to outbreaks in the northeast, which might be related to differences in the quality of reporting between these regions. For example, if surveillance in certain regions only captured larger, more easily detectable outbreaks with higher attack rates, this could bias our estimates of transmissibility upwards. Tremendous variability exists in outbreak reporting between states, ≥100-fold difference between the highest and lowest reporting states, which

Table 1. Norovirus outbreaks with exposed population size reported to the National Outbreak Reporting System, United States, 2009–2017*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. (%)</th>
<th>Median attack rate, % (IQR)</th>
<th>Median final size, % (IQR)</th>
<th>Median $R_0$ (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All outbreaks</td>
<td>7,094 (100)</td>
<td>22 (11–36)</td>
<td>28 (16–47)</td>
<td>2.75 (2.38–3.65)</td>
</tr>
<tr>
<td>Major setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child day care</td>
<td>272 (4)</td>
<td>21 (13–36)</td>
<td>18 (11–29)</td>
<td>2.67 (2.39–3.60)</td>
</tr>
<tr>
<td>Hospital or healthcare facility</td>
<td>271 (4)</td>
<td>22 (11–38)</td>
<td>19 (11–34)</td>
<td>2.70 (2.33–3.59)</td>
</tr>
<tr>
<td>Long-term care or assisted living facility</td>
<td>5,335 (75)</td>
<td>23 (13–36)</td>
<td>30 (17–47)</td>
<td>2.81 (2.42–3.76)</td>
</tr>
<tr>
<td>Other</td>
<td>350 (5)</td>
<td>20 (10–36)</td>
<td>24 (15–40)</td>
<td>2.66 (2.35–3.60)</td>
</tr>
<tr>
<td>Private home or residence</td>
<td>42 (1)</td>
<td>66 (50–91)</td>
<td>9 (6–16)</td>
<td>3.80 (2.26–4.92)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>77 (1)</td>
<td>50 (27–64)</td>
<td>10 (6–16)</td>
<td>3.12 (2.53–4.31)</td>
</tr>
<tr>
<td>School, college, or university</td>
<td>747 (11)</td>
<td>12 (6–24)</td>
<td>42 (19–80)</td>
<td>2.41 (2.24–2.92)</td>
</tr>
<tr>
<td>Season</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>4,016 (57)</td>
<td>22 (12–36)</td>
<td>30 (17–51)</td>
<td>2.80 (2.40–3.77)</td>
</tr>
<tr>
<td>Fall</td>
<td>808 (11)</td>
<td>21 (11–39)</td>
<td>26 (15–47)</td>
<td>2.72 (2.36–3.63)</td>
</tr>
<tr>
<td>Spring</td>
<td>1,964 (28)</td>
<td>20 (11–35)</td>
<td>27 (15–44)</td>
<td>2.69 (2.37–3.57)</td>
</tr>
<tr>
<td>Summer</td>
<td>306 (4)</td>
<td>17 (9–33)</td>
<td>19 (11–32)</td>
<td>2.57 (2.29–3.33)</td>
</tr>
<tr>
<td>Outbreak status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmed</td>
<td>3,114 (44)</td>
<td>26 (15–40)</td>
<td>35 (20–55)</td>
<td>2.99 (2.51–4.22)</td>
</tr>
<tr>
<td>Suspected</td>
<td>3,980 (56)</td>
<td>18 (9–31)</td>
<td>24 (14–40)</td>
<td>2.59 (2.32–3.27)</td>
</tr>
<tr>
<td>Census region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>1,898 (27)</td>
<td>17 (9–29)</td>
<td>31 (17–53)</td>
<td>2.58 (2.30–3.23)</td>
</tr>
<tr>
<td>Midwest</td>
<td>2,205 (31)</td>
<td>25 (13–39)</td>
<td>26 (15–44)</td>
<td>2.87 (2.44–3.98)</td>
</tr>
<tr>
<td>South</td>
<td>2,224 (31)</td>
<td>23 (12–38)</td>
<td>29 (17–47)</td>
<td>2.81 (2.39–3.93)</td>
</tr>
<tr>
<td>West</td>
<td>767 (11)</td>
<td>21 (13–34)</td>
<td>28 (16–44)</td>
<td>2.75 (2.42–3.57)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009 Jan–Jun†</td>
<td>243 (3)</td>
<td>28 (15–42)</td>
<td>35 (20–55)</td>
<td>3.09 (2.50–4.56)</td>
</tr>
<tr>
<td>2013 Jul–2014 Jun</td>
<td>913 (13)</td>
<td>20 (11–33)</td>
<td>29 (18–51)</td>
<td>2.68 (2.38–3.45)</td>
</tr>
<tr>
<td>2015 Jul–2016 Jun</td>
<td>1,007 (14)</td>
<td>17 (9–32)</td>
<td>25 (14–42)</td>
<td>2.57 (2.31–3.29)</td>
</tr>
</tbody>
</table>

*IQR, interquartile range.
†Partial norovirus years included in this analysis. The National Outbreak Reporting System was established in January 2009, and the first year of this analysis is 2009 January–June. At the time of analysis, we received data through December 2017.
likely affects the observed outbreak characteristics we included (34). Similarly, NORS has been collecting outbreak reports since January 2009, but in August 2012 CDC began a concerted effort to improve norovirus outbreak reporting to NORS and CaliciNet with the introduction of NoroSTAT (35,36). Thus, our finding that norovirus outbreaks reported before August 2012 were larger and had higher estimated \( R_0 \) and \( R_e \) values might be related to CDC’s efforts to capture outbreaks that previously would not have been reported, such as smaller outbreaks. Further, because the transmission mode can be difficult to identify for norovirus outbreaks, our analysis might have included outbreaks for which the mode of transmission was misclassified as person-to-person. Larger outbreaks with higher transmission are more likely to be reported, such as smaller outbreaks. In addition, the exposed population size is difficult to quantify and is not consistently reported, such as smaller outbreaks. Thus, the differences we found in estimated attack rates across different settings could be due to true variability in the exposed population size across settings or variability in the reliable reporting of the exposed population size. However, our analysis restricted to outbreaks in long-term care and assisted living facilities found the same trends among the variables for outbreak status, census region, season, and year as our analysis of all outbreaks, which suggests the results are robust.

Our study has several additional limitations. First, our process of data selection might have introduced bias into our analyses. We excluded outbreaks that occurred in multiple states, which are likely to have higher transmissibility given the larger geographic range involved; however, only 8 multistate outbreaks occurred during the study period, thus the bias is likely negligible. A substantial proportion of the dataset, 5,573 (31%) outbreaks, had to be excluded because the exposed population size was not reported. Excluding these outbreaks could introduce bias if the exposed population size is more likely to be reported for outbreaks with smaller, or larger, exposed population sizes. We only included outbreaks with person-to-person transmission; thus, our estimates of transmissibility are not generalizable to norovirus outbreaks where transmission occurs via other modes, such as foodborne, waterborne, or environmental transmission.

A second set of limitations relates to the final size method. This method assumes a susceptible-infected-recovered type infection in a homogeneously mixing population (12), but this simplification likely does not reflect true mixing patterns. In addition, we might observe different mixing patterns in each of the different outbreak settings, such as older persons in long-term care facilities versus young children in childcare. The final size method also underestimates reproduction numbers for outbreaks with high attack rates. For example, in private homes, attack rates are high, but exposed population sizes are small. If everyone in the household is infected, then no additional infections can occur in the home. Thus, the final size method cannot capture any additional transmission that could have happened if the exposed population size had been larger, such as a higher number of persons in the household. Becker termed this limitation the “wasted infection potential” (37). Further, the final size method does not account for the effect of control measures. For some of the outbreaks represented in our dataset, control measures were most likely implemented, such as isolating ill persons and cleaning contamination. Such interventions likely would reduce the number ill, and the estimated \( R_0 \) would be lower than the \( R_0 \) in the absence of control measures.

### Table 2. Estimated log-linear change in \( R_0 \) from the intercept for linear regression of log transformed \( R_0 \) for norovirus outbreaks reported to the National Outbreak Reporting System, United States, 2009–2017

<table>
<thead>
<tr>
<th>Category</th>
<th>Estimated log-linear change in ( R_0 ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.35 (3.26–3.45)</td>
</tr>
<tr>
<td>Major setting</td>
<td></td>
</tr>
<tr>
<td>Long-term care or assisted living facility</td>
<td>Referent</td>
</tr>
<tr>
<td>Child care</td>
<td>0.99 (0.95–1.03)</td>
</tr>
<tr>
<td>Hospital or healthcare facility</td>
<td>0.93 (0.90–0.97)</td>
</tr>
<tr>
<td>Other</td>
<td>0.97 (0.93–1.01)</td>
</tr>
<tr>
<td>Private home or residence</td>
<td>0.99 (0.82–1.19)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>1.01 (0.91–1.11)</td>
</tr>
<tr>
<td>School, college, or university</td>
<td>0.87 (0.85–0.89)</td>
</tr>
<tr>
<td>Season</td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>Referent</td>
</tr>
<tr>
<td>Fall</td>
<td>1.00 (0.98–1.03)</td>
</tr>
<tr>
<td>Spring</td>
<td>0.98 (0.96–1.00)</td>
</tr>
<tr>
<td>Summer</td>
<td>0.93 (0.89–0.96)</td>
</tr>
<tr>
<td>Outbreak status</td>
<td></td>
</tr>
<tr>
<td>Confirmed</td>
<td>Referent</td>
</tr>
<tr>
<td>Suspected</td>
<td>0.90 (0.88–0.92)</td>
</tr>
<tr>
<td>Census region</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>Referent</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.89 (0.87–0.91)</td>
</tr>
<tr>
<td>Midwest</td>
<td>1.00 (0.97–1.02)</td>
</tr>
<tr>
<td>West</td>
<td>0.98 (0.95–1.01)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
</tr>
<tr>
<td>2009 Jan–Jun</td>
<td>1.16 (1.10–1.23)</td>
</tr>
<tr>
<td>2009 Jul–2010 Jun</td>
<td>1.17 (1.11–1.23)</td>
</tr>
<tr>
<td>2010 Jul–2011 Jun</td>
<td>1.16 (1.12–1.21)</td>
</tr>
<tr>
<td>2011 Jul–2012 Jun</td>
<td>1.12 (1.08–1.16)</td>
</tr>
<tr>
<td>2012 Jul–2013 Jun</td>
<td>1.04 (1.01–1.07)</td>
</tr>
<tr>
<td>2013 Jul–2014 Jun</td>
<td>1.02 (0.99–1.06)</td>
</tr>
<tr>
<td>2014 Jul–2015 Jun</td>
<td>1.05 (1.02–1.08)</td>
</tr>
<tr>
<td>2015 Jul–2016 Jun</td>
<td>1.02 (0.99–1.05)</td>
</tr>
<tr>
<td>2016 Jul–2017 Jun</td>
<td>Referent</td>
</tr>
<tr>
<td>2017 Jul–Dec</td>
<td>1.04 (1.00–1.09)</td>
</tr>
</tbody>
</table>
In addition, the final size method assumes that the proportion of susceptible persons is known at the start of an outbreak; however, the level of susceptibility to norovirus is not well known. Certain host genetic factors are associated with the ability of norovirus to establish an infection within a human host (38–41), leading to variable susceptibility to norovirus infection (42–44). Secretor-negative persons have nonfunctional fucosyltransferase-2 genes, causing infection failure for norovirus genogroups I and II type 4 (38,40,41,45,46). Our estimates of \( R_0 \) and \( R_e \) assume that 47% of the population in our dataset is susceptible at the start of all outbreaks. However, the proportion susceptible varies among outbreaks and potentially over time and age as the distribution of circulating norovirus genotypes change. Further, our regression model estimates were sensitive to our assumption of the percent susceptible at the start of an outbreak. When we assumed 47% and 80% of the population was susceptible, the estimated transmissibility of norovirus in private homes or residences and restaurants was higher than transmissibility in long-term care and assisted living facilities. However, when we assumed 27% of the population was susceptible at the start of an outbreak, the association between private homes or residences and restaurants reversed. These settings then had lower estimated transmission relative to outbreaks in long-term care and assisted living facilities because the population size that can be infected is much lower, thus reducing the estimates of \( R_0 \) and \( R_e \). For example, if a household had 15 persons, the maximum possible \( R_e \) assuming 27% susceptibility is 4, which is lower than the average predicted \( R_e \) for outbreaks in the reference group. Therefore, the results for private homes or residences and restaurants, where exposed population sizes are lower, should be interpreted with caution because transmission values in these settings might be underestimated. We also assumed that only symptomatic persons contribute to transmission in our calculation; persons with asymptomatic norovirus infections can contribute to transmission, but they likely are not as infectious as persons with symptomatic infections (22,47).

Finally, our main analysis does not account for norovirus genotype. Because of the limited data available on outbreak genotype we were not able to fully assess whether certain genotypes were more transmissible. As more genotyping data become available, future studies should investigate transmissibility.

We estimated reproduction numbers by using the final size method for >7,000 outbreaks from a national outbreak reporting system, then used these estimates to examine factors associated with norovirus transmission. Our analyses suggest that norovirus transmission rates are modest. Such modest rates of \( R_0 \) suggest there are opportunities for effective control measures to curtail transmission of norovirus. However, challenges remain. Transmission by asymptomatic persons, which we did not account for in this analysis and generally goes undetected in surveillance, can limit the effectiveness of traditional control methods focused on ill persons, even for pathogens with modest transmission (48).

Overall, we found limited variation in \( R_0 \) and \( R_e \) for reported norovirus outbreaks in the United States, particularly across different settings. Our findings highlight the need for better data on the total exposed population sizes in outbreaks, which heavily influence estimates of attack rates, \( R_{sp} \) and \( R_{rec} \) to further refine estimates of these outbreak factors.

This work was supported by funding from the Norovirus Vaccine Modeling (NoVaMod) project (grant no. R01 GM124280), the National Institutes of Health, Agency for Healthcare Research Quality (grant no. R01 HS025987), and the National Institutes of Health graduate and postdoctoral training grant in environmental health science and toxicology (grant no. T32 ES012870). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the US Department of Health and Human Services.

About the Author

Dr. Steele is an epidemiologist in the Influenza Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention. Her research interests include infectious disease dynamics and predicting public health impacts of vaccination.

References


2. Wikswo ME, Kambhampati A, Shioda K, Walsh KA, Bowen AB, Hall AJ. Outbreaks of acute gastroenteritis transmitted by person-to-person contact, environmental contamination, and unknown modes of transmission–United States, 2009–2013. MMWR Surveill Summ. 2015;64:1–16 https://doi.org/10.15585/mmwr.mm6412a1


Characterizing Norovirus Transmission

Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 26, No. 8, August 2020

Address for correspondence: Molly Steele, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, Mailstop H24-7; Atlanta, GA 30329-4027, USA; email: MSteele@cdc.gov
Characterizing Norovirus Transmission from Outbreak Data, United States

Appendix

Estimating Basic and Effective Reproduction Numbers

We calculated basic reproduction ($R_0$), effective reproduction ($R_e$), and the associated standard errors (SE) for each outbreak. We used equations proposed by Becker (1) that use the final epidemic size

$$R_0 = \frac{N - 1}{C} \sum_{i=S-C+1}^{S} \frac{1}{i}$$

$$SE(R_0) = \frac{N - 1}{C} \sqrt{\sum_{i=S-C+1}^{S} \frac{1}{i^2} + \frac{CR_0^2}{(N - 1)^2}}$$

where $N$ is the total population size, $C$ is the total number of cases in the outbreak, and $S$ is the number of susceptible persons at the start of the outbreak. $R_e$ is calculated by replacing $N$ with $S$ in the first equation above. $SE(R_e)$ is calculated by replacing $N$ with $S$ and $R_0$ with $R_e$ in the second equation shown above. The final size method assumes a susceptible-infected-recovered type infection with a closed, homogenously mixing population (1).

Genogroup and genotype analysis

Among the outbreaks for which we calculated $R_0$ and $R_e$ only 1,571 outbreaks (22%) had data on norovirus genotype. We fit a linear regression model to the log transformed estimate of $R_0$ to assess whether the following genogroup and genotypes were associated transmissibility: genogroup I (GI), genogroup II genotype 4 (GII.4) and genogroup II non-genotype 4 (GII.non4). We found that $R_0$ varied little for outbreaks of GI ($R_0 = 3.50$ [95% CI 3.32, 3.68]), GII.4 ($R_0 = 3.46$ [95% CI 3.38, 3.55]) and GII.non4 ($R_0 = 3.26$ [95% CI 3.11, 3.41]).
genotype and year to assess whether GII.4 noroviruses had different effects on transmissibility at the time when new strains emerged (i.e., GII.4 New Orleans in the 2009/2010 season and GII.4 Sydney in the 2012/2013 season). There was not sufficient evidence to suggest that $R_0$ differed for outbreaks of GII.4 during the norovirus seasons when GII.4 New Orleans ($R_0 = 4.00$ [95% CI 3.53, 4.54]) or GII.4 Sydney ($R_0 = 3.38$ [95% CI 3.24, 3.53]) relative to GII.4 outbreaks reported in seasons when no new strains emerged ($R_0 = 3.38$ [95% CI 3.19, 3.58]). (Appendix Table 4)

**Alternative Models**

We assessed two alternative approaches for modeling norovirus transmission: a logistic regression to model a binary transmission outcome (i.e., high versus low transmission) and a negative binomial regression to model the final size of outbreaks, adjusting for exposed population size (i.e., modeling attack rates). For our logistic regression, we used the first and third tertiles of estimated values of $R_0$ and $R_e$, assuming the percent susceptible was 47%, to determine the cutoffs for our outcome of interest: low versus high transmission. We excluded outbreaks with transmission values within the second tertile and focus our logistic regression comparison between the lowest and highest tertiles of transmission. The third tertile of $R_0$ and $R_e$ values were 3.23 and 1.52, respectively.

The trends of transmissibility across our variables of interest (outbreak setting, census region, season, year, whether norovirus was suspected or confirmed and norovirus genotype) from our main regression analysis of a continuous transmission outcome were consistent across the logistic regressions of high $R_0$ ($R_0>3.23$) and $R_e$ ($R_e>1.52$) and linear regression of $R_e$ values. (Appendix Table 5) The trends of transmissibility were consistent for most of our variables of interest in the negative binomial model of final outbreak sizes; however, private homes or residences and restaurants had a much more pronounced effect on the attack rate of outbreaks, relative to long-term care and assisted living facilities ($RR = 2.35$ (95% CI 1.85, 3.01) and $RR = 1.67$ (95% CI 1.40, 2.01), respectively). As the exposed population size is difficult to quantify, and thus may not be reported reliably, we analyzed the subset of outbreaks that occurred within long-term care and assisted living facilities with our regression models. The patterns found among the variables for outbreak status, census region, season and year were consistent with what was found analyzing the full dataset. (Appendix Table 6)
Percent Susceptible

We tested the sensitivity of our results to the assumption of the percent susceptible start of an outbreak by running all the regression models assuming the percent susceptible at the start was 27% and 80%, which represent the minimum and maximum estimates of the percent susceptible to AGE from published challenge studies, respectively (Appendix Table 2). By adjusting our assumption of the percent susceptible at the start of the outbreak to 27% and 80%, the median $R_0$ was 6.04 (IQR 4.53, 9.38) and 1.43 (IQR 1.33, 1.61), respectively, while the median $R_e$ was 1.82 (IQR 1.24, 3.83) and 1.14 (IQR 1.07, 1.29), respectively. Assuming the proportion susceptible was 27% we found that outbreaks in long-term care and assisted living facilities were more likely to have $R_0 > 8.05$ and $R_e > 3.24$ relative to all other settings. (Appendix Table 5, Appendix Table 6, Appendix Figure 2) When the percent susceptible was 80% we found that outbreaks in long-term care and assisted living facilities had increased odds of having $R_0 > 1.54$ and $R_e > 1.23$ compared to hospitals/other healthcare facilities, schools/colleges/universities, and other settings. Outbreaks in private homes or residences and restaurants had higher odds of having $R_0 > 1.54$ and $R_e > 1.23$ relative to long-term care and assisted living facilities, however the confidence intervals are wide due to small sample sizes. (Appendix Table 7, Appendix Table 8, Appendix Figure 2) Trends in the variables for census region, season, year, and whether norovirus was suspected or confirmed for the models assuming 27% and 80% susceptibility were consistent with the models assuming 47% susceptibility. (Table 2, Appendix Table 7, Appendix Table 8)

### Appendix Table 1. Descriptive characteristics of norovirus outbreaks that met our inclusion criteria and those outbreaks that were excluded from this analysis

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. (%)</th>
<th>Median attack rate (IQR)*</th>
<th>Median final size (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Outbreaks</td>
<td>10,728 (60)</td>
<td>22 (10–44)</td>
<td>20 (10–39)</td>
</tr>
<tr>
<td>Major setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child day care</td>
<td>168 (2)</td>
<td>21 (16–36)</td>
<td>16 (10–26)</td>
</tr>
<tr>
<td>Hospital/healthcare facility</td>
<td>316 (3)</td>
<td>24 (9–54)</td>
<td>16 (6–28)</td>
</tr>
<tr>
<td>Long-term care/assisted living facility</td>
<td>3,596 (34)</td>
<td>21 (10–42)</td>
<td>25 (14–43)</td>
</tr>
<tr>
<td>Other</td>
<td>493 (5)</td>
<td>24 (11–42)</td>
<td>19 (9–40)</td>
</tr>
<tr>
<td>Private home/residence</td>
<td>81 (1)</td>
<td>61 (40–95)</td>
<td>8 (5–13)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>259 (2)</td>
<td>41 (17–88)</td>
<td>8 (4–13)</td>
</tr>
<tr>
<td>School/college/university</td>
<td>562 (5)</td>
<td>12 (6–26)</td>
<td>30 (13–69)</td>
</tr>
<tr>
<td>Missing</td>
<td>5,253 (49)</td>
<td>30 (14–43)</td>
<td>18 (6–36)</td>
</tr>
<tr>
<td>Season</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>5,597 (49)</td>
<td>22 (10–4)</td>
<td>23 (11–44)</td>
</tr>
<tr>
<td>Fall</td>
<td>1,390 (13)</td>
<td>20 (10–40)</td>
<td>18 (8–35)</td>
</tr>
<tr>
<td>Spring</td>
<td>3,137 (30)</td>
<td>21 (10–43)</td>
<td>19 (9–36)</td>
</tr>
<tr>
<td>Summer</td>
<td>868 (8)</td>
<td>23 (10–41)</td>
<td>14 (7–26)</td>
</tr>
<tr>
<td>Outbreak Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmed</td>
<td>4,875 (55)</td>
<td>22 (11–42)</td>
<td>23 (11–42)</td>
</tr>
<tr>
<td>Suspected</td>
<td>5,853 (45)</td>
<td>21 (9–44)</td>
<td>18 (9–36)</td>
</tr>
<tr>
<td>Census region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>2,667 (25)</td>
<td>11 (6–24)</td>
<td>24 (12–47)</td>
</tr>
</tbody>
</table>
The selection process for a linear regression model of log-transformed R0 values*

The number of challenged participants who subsequently developed acute gastroenteritis*

Data from published norovirus challenge studies on the number of participants challenged with norovirus and the number of challenged participants who subsequently developed acute gastroenteritis*

### Appendix Table 2. Data from published norovirus challenge studies on the number of participants challenged with norovirus and the number of challenged participants who subsequently developed acute gastroenteritis*

<table>
<thead>
<tr>
<th>Study (reference)</th>
<th>Genogroup and genotype</th>
<th>No. challenged</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolin 1972 (2)</td>
<td>GI.4</td>
<td>1550 (14)</td>
<td>23 (11–38)</td>
</tr>
<tr>
<td>Wyatt 1974 (3)</td>
<td>GII.4</td>
<td>774 (7)</td>
<td>19 (9–34)</td>
</tr>
<tr>
<td>Parrino 1977 (4)</td>
<td>GII.4</td>
<td>436 (4)</td>
<td>21 (10–40)</td>
</tr>
<tr>
<td>Treanor 1988 (5)</td>
<td>GII.4</td>
<td>7,968 (74)</td>
<td>20 (9–41)</td>
</tr>
<tr>
<td>Johnson 1990 (6)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham 1994 (7)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindesmith 2003 (8)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindesmith 2005 (9)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmar 2008 (10)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leon 2011 (11)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmar 2011 (12)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seitz 2011 (13)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French 2012 (14)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bernstein 2015 (15)</td>
<td>GII.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The medians and IQRs for attack rates were calculated for the subset of outbreaks where both the exposed population size and total estimated primary cases were reported. Among the 10,726 outbreaks excluded from our primary analysis, 8,903 outbreaks were missing data for the exposed population size.
- Two partial norovirus years included in this analysis. NORS was established in January 2009, thus the first year of this analysis is January–June 2009. At the time of analysis, we received data through December 2017, thus the final year of this analysis is July–Dec 2017.

### Appendix Table 3. Estimated log linear change in R0 (95% CI) from the estimated R0 for the intercept for each model in a forward selection process for a linear regression model of log-transformed R0 values*

<table>
<thead>
<tr>
<th>Model variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.22 (3.19–3.25)</td>
<td>3.44 (3.39–3.49)</td>
<td>3.57 (3.50–3.64)</td>
<td>3.61 (3.53–3.68)</td>
<td>3.35 (3.26–3.45)</td>
</tr>
<tr>
<td>Child day care</td>
<td>0.95 (0.91–0.99)</td>
<td>0.99 (0.95–1.03)</td>
<td>0.98 (0.94–1.02)</td>
<td>0.98 (0.94–1.02)</td>
<td>0.99 (0.95–1.03)</td>
</tr>
<tr>
<td>Hospital or healthcare facility</td>
<td>0.94 (0.90–0.98)</td>
<td>0.93 (0.89–0.97)</td>
<td>0.94 (0.90–0.98)</td>
<td>0.94 (0.90–0.98)</td>
<td>0.93 (0.90–0.97)</td>
</tr>
<tr>
<td>Other</td>
<td>0.96 (0.92–1.00)</td>
<td>0.96 (0.93–1.00)</td>
<td>0.95 (0.91–0.99)</td>
<td>0.97 (0.93–1.01)</td>
<td>0.97 (0.93–1.01)</td>
</tr>
<tr>
<td>Private home or residence</td>
<td>1.00 (0.84–1.19)</td>
<td>0.99 (0.83–1.18)</td>
<td>0.98 (0.81–1.17)</td>
<td>0.99 (0.82–1.19)</td>
<td>0.99 (0.82–1.19)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>1.02 (0.93–1.12)</td>
<td>1.01 (0.92–1.11)</td>
<td>1.00 (0.91–1.09)</td>
<td>1.00 (0.91–1.10)</td>
<td>1.01 (0.91–1.11)</td>
</tr>
<tr>
<td>School, college, or university</td>
<td>0.84 (0.82–0.87)</td>
<td>0.88 (0.86–0.90)</td>
<td>0.86 (0.84–0.88)</td>
<td>0.86 (0.84–0.88)</td>
<td>0.87 (0.85–0.89)</td>
</tr>
<tr>
<td>Suspected outbreak</td>
<td>-</td>
<td>0.88 (0.87–0.90)</td>
<td>0.88 (0.87–0.90)</td>
<td>0.89 (0.87–0.90)</td>
<td>0.90 (0.88–0.92)</td>
</tr>
<tr>
<td>Region 1</td>
<td>-</td>
<td>-</td>
<td>0.89 (0.87–0.91)</td>
<td>0.89 (0.87–0.91)</td>
<td>0.89 (0.87–0.91)</td>
</tr>
<tr>
<td>Region 2</td>
<td>-</td>
<td>-</td>
<td>0.99 (0.97–1.02)</td>
<td>0.99 (0.97–1.02)</td>
<td>1.00 (0.97–1.02)</td>
</tr>
<tr>
<td>Region 4</td>
<td>-</td>
<td>-</td>
<td>0.96 (0.93–0.99)</td>
<td>0.96 (0.93–0.99)</td>
<td>0.98 (0.95–1.01)</td>
</tr>
<tr>
<td>Fall</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00 (0.97–1.03)</td>
<td>1.00 (0.98–1.03)</td>
</tr>
<tr>
<td>Spring</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.97 (0.95–0.99)</td>
<td>0.98 (0.96–1.00)</td>
</tr>
<tr>
<td>Summer</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.92 (0.88–0.95)</td>
<td>0.93 (0.89–0.96)</td>
</tr>
</tbody>
</table>

*We assume that the average proportion that develop AGE across all studies, weighted by total number of participants in each study, is the proportion that are susceptible to norovirus in our calculations of R0 and Re.

†Average proportion susceptible weighted by number of participants.
Table 5. Risk ratios of attack rates, estimated log-linear change in $R_0$ and $R_e$ relative to the intercept from linear regression of the log transformed reproduction numbers and odds ratios of an outbreak with high transmission from logistic regression.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>RR of attack rates (95% CI)</th>
<th>OR of $R_0$&gt;3.23 (95% CI)</th>
<th>Estimated log-linear change in $R_0$ (95% CI)</th>
<th>OR of $R_e$&gt;1.52 (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.27 (0.26–0.29)</td>
<td>1.75 (1.41–2.18)</td>
<td>1.63 (1.58–1.68)</td>
<td>1.72 (1.39–2.13)</td>
</tr>
<tr>
<td>Major setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term care, assisted living</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>facility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child day care</td>
<td>1.08 (0.98–1.19)</td>
<td>1.03 (0.74–1.42)</td>
<td>1.00 (0.95–1.05)</td>
<td>1.11 (0.80–1.53)</td>
</tr>
<tr>
<td>Hospital or healthcare facility</td>
<td>0.94 (0.86–1.03)</td>
<td>0.62 (0.45–0.85)</td>
<td>0.94 (0.89–0.98)</td>
<td>0.65 (0.48–0.89)</td>
</tr>
<tr>
<td>Other</td>
<td>0.98 (0.90–1.06)</td>
<td>0.73 (0.55–0.96)</td>
<td>0.98 (0.93–1.02)</td>
<td>0.74 (0.56–0.98)</td>
</tr>
<tr>
<td>Private home or residence</td>
<td>2.35 (1.85–3.01)</td>
<td>1.80 (0.88–3.87)</td>
<td>1.31 (1.14–1.51)</td>
<td>8.47 (3.22–29.27)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>1.67 (1.40–2.01)</td>
<td>1.41 (0.78–2.62)</td>
<td>1.10 (0.99–1.22)</td>
<td>1.96 (1.09–3.67)</td>
</tr>
<tr>
<td>School, college, or university</td>
<td>0.67 (0.63–0.71)</td>
<td>0.29 (0.23–0.36)</td>
<td>0.86 (0.83–0.88)</td>
<td>0.30 (0.24–0.38)</td>
</tr>
<tr>
<td>Season</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Fall</td>
<td>0.99 (0.94–1.05)</td>
<td>1.00 (0.81–1.23)</td>
<td>1.01 (0.97–1.04)</td>
<td>0.97 (0.79–1.19)</td>
</tr>
<tr>
<td>Spring</td>
<td>0.97 (0.93–1.01)</td>
<td>0.90 (0.78–1.04)</td>
<td>0.98 (0.95–1.00)</td>
<td>0.89 (0.77–1.04)</td>
</tr>
<tr>
<td>Summer</td>
<td>0.86 (0.79–0.95)</td>
<td>0.65 (0.47–0.88)</td>
<td>0.92 (0.88–0.96)</td>
<td>0.62 (0.45–0.85)</td>
</tr>
<tr>
<td>Outbreak Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmed outbreak</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Suspected outbreak</td>
<td>0.82 (0.79–0.85)</td>
<td>0.43 (0.37–0.49)</td>
<td>0.89 (0.87–0.91)</td>
<td>0.42 (0.37–0.47)</td>
</tr>
<tr>
<td>Census region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.77 (0.74–0.81)</td>
<td>0.44 (0.37–0.52)</td>
<td>0.88 (0.85–0.90)</td>
<td>0.45 (0.38–0.53)</td>
</tr>
<tr>
<td>Midwest</td>
<td>1.07 (1.02–1.11)</td>
<td>1.02 (0.87–1.20)</td>
<td>1.00 (0.98–1.03)</td>
<td>1.03 (0.88–1.20)</td>
</tr>
<tr>
<td>West</td>
<td>0.99 (0.93–1.05)</td>
<td>0.95 (0.76–1.19)</td>
<td>0.97 (0.93–1.00)</td>
<td>0.94 (0.75–1.17)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 2009–Jun 2009</td>
<td>1.34 (1.21–1.48)</td>
<td>2.49 (1.73–3.62)</td>
<td>1.20 (1.12–1.28)</td>
<td>2.50 (1.73–3.63)</td>
</tr>
<tr>
<td>Jul 2009–Jun 2010</td>
<td>1.37 (1.25–1.52)</td>
<td>2.59 (1.82–3.73)</td>
<td>1.22 (1.15–1.30)</td>
<td>2.55 (1.79–3.65)</td>
</tr>
<tr>
<td>Jul 2010–Jun 2011</td>
<td>1.34 (1.25–1.45)</td>
<td>2.64 (2.01–3.49)</td>
<td>1.20 (1.14–1.25)</td>
<td>2.58 (1.97–3.39)</td>
</tr>
<tr>
<td>Jul 2011–Jun 2012</td>
<td>1.24 (1.15–1.33)</td>
<td>2.12 (1.63–2.77)</td>
<td>1.14 (1.09–1.19)</td>
<td>2.16 (1.66–2.81)</td>
</tr>
</tbody>
</table>
### Appendix Table 6. Estimated log-linear change in basic and effective reproduction numbers relative to the intercept from linear regression*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Basic reproduction number</th>
<th>Effective reproduction number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR of attack rates (95% CI)</td>
<td>Estimated log-linear change (95% CI)</td>
</tr>
<tr>
<td>Jul 2012–Dec 2017</td>
<td>1.05 (0.96–1.14)</td>
<td>1.40 (1.03–1.90)</td>
</tr>
</tbody>
</table>

*Assuming the percent susceptible at the start of an outbreak is 47%.
†Logistic regression compares outbreaks with transmission in the third tertile (Re>1.52) to outbreaks in the first tertile (Re<1.17) and does not include Re values in second tertile. Linear and negative binomial regressions use full dataset.

### Appendix Table 7. Estimated log-linear change in basic reproductive number (R<sub>0</sub>) from linear regression and odds ratios of outbreaks with high R<sub>0</sub> assuming 27% and 80% susceptible at the start of the outbreak*  

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>27% Susceptible</th>
<th>80% Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated log-linear change in R&lt;sub&gt;0&lt;/sub&gt; (95% CI)</td>
<td>OR of R&lt;sub&gt;0&lt;/sub&gt;&gt;8.05 (95% CI)†</td>
</tr>
<tr>
<td>Intercept</td>
<td>7.42 (7.18–7.68)</td>
<td>2.05 (1.66–2.54)</td>
</tr>
</tbody>
</table>

*Odds ratios of outbreaks with R<sub>0</sub>&gt;3.23 and R<sub>0</sub>&gt;1.52, final size adjusting for exposed population size among long-term care or assisted care facilities, assuming 47% susceptible at the start of the outbreak. R<sub>0</sub> estimated reproduction number.
†Logistic regression compares outbreaks with transmission in the third tertile (R<sub>0</sub>&gt;3.23, R<sub>0</sub>&gt;1.52) to outbreaks in the first tertile (R<sub>0</sub>&lt;2.48, R<sub>0</sub>&lt;1.17), and does not include R<sub>0</sub> values in second tertile. Linear and negative binomial regressions use full dataset.
### Appendix Table 8. Estimated log-linear change in $R_0$ (95% CI) from linear regression and odds ratios of outbreaks with high $R_0$ assuming 27% and 80% susceptible at the start of the outbreak

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>27% Susceptible</th>
<th>80% Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated log-linear change in $R_0$ (95% CI)</td>
<td>OR of $R_0&gt;3.24$ (95% CI)</td>
</tr>
<tr>
<td>Probable outbreak</td>
<td>0.37 (0.95–0.99)</td>
<td>0.45 (0.40–0.51)</td>
</tr>
<tr>
<td>Census Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.93 (0.91–0.96)</td>
<td>0.55 (0.47–0.65)</td>
</tr>
<tr>
<td>Midwest</td>
<td>0.99 (0.97–1.02)</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>West</td>
<td>1.03 (0.99–1.06)</td>
<td>1.10 (0.88–1.37)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 2009–Jun 2009</td>
<td>1.12 (1.06–1.20)</td>
<td>2.05 (1.42–3.00)</td>
</tr>
<tr>
<td>Jul 2009–Jun 2010</td>
<td>1.11 (1.05–1.18)</td>
<td>1.86 (1.30–2.67)</td>
</tr>
<tr>
<td>Jul 2010–Jun 2011</td>
<td>1.11 (1.07–1.16)</td>
<td>1.95 (1.49–2.57)</td>
</tr>
<tr>
<td>Jul 2011–Jun 2012</td>
<td>1.13 (1.08–1.18)</td>
<td>2.09 (1.62–2.72)</td>
</tr>
<tr>
<td>Jul 2012–Jun 2013</td>
<td>1.02 (0.99–1.06)</td>
<td>1.13 (0.90–1.41)</td>
</tr>
<tr>
<td>Jul 2013–Jun 2014</td>
<td>1.01 (0.98–1.05)</td>
<td>1.12 (0.88–1.42)</td>
</tr>
<tr>
<td>Jul 2014–Jun 2015</td>
<td>1.03 (0.99–1.07)</td>
<td>1.26 (1.00–1.59)</td>
</tr>
<tr>
<td>Jul 2015–Jun 2016</td>
<td>1.00 (0.96–1.03)</td>
<td>1.01 (0.80–1.26)</td>
</tr>
<tr>
<td>Jul 2016–Jun 2017</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Jul 2017–Dec 2017</td>
<td>1.04 (0.99–1.09)</td>
<td>1.22 (0.91–1.65)</td>
</tr>
</tbody>
</table>

*OR, odds ratio; $R_0$, basic reproduction number; $R_0^*$, estimated reproduction number.
†Logistic regression compares outbreaks with transmission in the third tertile ($S = 27\%$: $R_0>4.96$; $S = 80\%$: $R_0<1.36$) to outbreaks in the first tertile ($S = 27\%$: $R_0<4.96$; $S = 80\%$: $R_0>1.36$) and does not include $R_0/R_0^*$ values in second tertile. Linear and negative binomial regressions use full dataset.
Appendix Figure 1. Flowchart of outbreaks included in analysis.
Appendix Figure 2. Reproduction rate, and estimated reproduction rate for norovirus in the United States assuming the initial proportion susceptible is 47%. Blue circles indicate estimated reproductive rate. Green circles indicate basic reproduction rate. $R_0$, basic reproduction rate; $R_e$, estimated reproduction rate.
Appendix Figure 3. Estimated $R_0$ (green squares) and $R_0$ (blue circles) assuming A) 27% susceptible and B) 80% susceptible at the start of the outbreak.

References


https://doi.org/10.1056/NEJM197707142970204


