The duration of immunity to norovirus (NoV) gastroenteritis has been believed to be from 6 months to 2 years. However, several observations are inconsistent with this short period. To gain better estimates of the duration of immunity to NoV, we developed a mathematical model of community NoV transmission. The model was parameterized from the literature and also fit to age-specific incidence data from England and Wales by using maximum likelihood. We developed several scenarios to determine the effect of unknowns regarding transmission and immunity on estimates of the duration of immunity. In the various models, duration of immunity to NoV gastroenteritis was estimated at 4.1 (95% CI 3.2–5.1) to 8.7 (95% CI 6.8–11.3) years. Moreover, we calculated that children (<5 years) are much more infectious than older children and adults. If a vaccine can achieve protection for duration of natural immunity indicated by our results, its potential health and economic benefits could be substantial.

Noroviruses (NoVs) are the most common cause of acute gastroenteritis (AGE) in industrialized countries. In the United States, NoV causes an estimated 21 million cases of AGE (1), 1.7 million outpatient visits (2), 400,000 emergency care visits, 70,000 hospitalizations (3), and 800 deaths annually across all age groups (4). Although the highest rates of disease are in young children, infection and disease occur throughout life (5), despite an antibody seroprevalence >50%, and infection rates approach 100% in older adults (6,7).

Frequently cited estimates of the duration of immunity to NoV are based on human challenge studies conducted in the 1970s. In the first, Parrino et al. challenged volunteers with Norwalk virus (the prototype NoV strain) inoculum multiple times. Results suggested that the immunity to Norwalk AGE lasts from ≈2 months to 2 years (8). A subsequent study with a shorter challenge interval suggested that immunity to Norwalk virus lasts for at least 6 months (9). In addition, the collection of volunteer studies together demonstrate that antibodies against NoV may not confer protection and that protection from infection (serologic response or viral shedding) is harder to achieve than protection from disease (defined as AGE symptoms) (10–14). That said, most recent studies have reported some protection from illness and infection in association with antibodies that block binding of virus-like particles to histo-blood group antigen (HBGA) (13,14). Other studies have also associated genetic resistance to NoV infections with mutations in the 1,2-fucosyltransferase (FUT2) gene (or “secretor” gene) (15). Persons with a nonsecretor gene (FUT2−/−) represent as much as 20% of the European population. Challenge studies have also shown that recently infected volunteers are susceptible to heterologous strains sooner than to homotypic challenge, indicating limited cross-protection (11).

One of many concerns with all classic challenge studies is that the virus dose given to volunteers was several thousand–fold greater than the small amount of virus capable of causing human illness (estimated as 18–1,000 virus particles) (16). Thus, immunity to a lower challenge dose, similar to what might be encountered in the community, might be more robust and broadly protective than the protection against artificial doses encountered in these volunteer studies. Indeed, Teunis et al. have clearly demonstrated a dose–response relationship whereby persons challenged with a higher NoV dose have substantially greater illness risk (16).

Furthermore, in contrast with results of early challenge studies, several observations can be made that, when taken together, are inconsistent with a duration of immunity on the scale of months. First, the incidence of NoV in the general population has been estimated in several countries as ≈5% per year, with substantially higher rates in children (5). Second, Norwalk virus (GI.1) volunteer studies conducted over 3 decades, indicate that approximately one third of genetically susceptible

Author affiliations: Rollins School of Public Health, Emory University, Atlanta, Georgia, USA (K. Simmons, J. Leon, B. Lopman); Centers for Disease Control and Prevention, Atlanta (K. Simmons, B. Lopman); and Imperial College London, London, UK (M. Gambhir)

DOI: http://dx.doi.org/10.3201/eid1908.130472

Kirsten Simmons, Manoj Gambhir, Juan Leon, and Ben Lopman
persons (i.e., secretor-positive persons with a functional FUT2 gene) are immune (Table 1) (18,20,22). The point prevalence of immunity in the population (i.e., population immunity) can be approximated by the incidence of infection (or exposure) multiplied by the duration of immunity. If duration of immunity is truly <1 year and incidence is 5%, <5% of the population should have acquired immunity at any given time. However, challenge studies show population immunity levels on the order of 30%–45%, suggesting that our understanding of the duration of immunity at any given time.

If duration of immunity is truly <1 year and incidence is related to NoV transmission remain unknown (e.g., the role asymptomatic persons who shed virus play in transmission). Therefore, we constructed and fit a series of 6 models to represent the variety of possible infection processes to gain a more robust estimate of the duration of immunity. This approach does not consider multiple strains or the emergence of new variants, so we are effectively estimating minimum duration of immunity in the absence of major strain changes.

Methods

Model Design

We developed a deterministic dynamic transmission model with age structure that tracks the population with respect to NoV infection and immunity status (Figure 1; online Technical Appendix, wwwnc.cdc.gov/EID/article/19/8/13-0472-Techapp1.pdf for model equations). Here we describe the basic structure of the model (model A), which forms the basis for 5 other iterations (models B–F, described below). The models track 5 classes of persons: 1) susceptible to infection and disease (S), 2) exposed but not yet symptomatic (E), 3) infected with symptoms (I), 4) infected but asymptomatic (A), and 5) immune to disease, but not infection (R). In model D, we included an additional class for genetically resistant persons (G).

We assume that maternal immunity is negligible because the youngest age class includes children ages 0–4 years; as such, newborns in all models except model D enter directly into S class. In model D, genetically resistant persons bypass the S class and remain resistant for life, although they make contacts and are included in calculations of incidence for model-fitting purposes, because all persons (not just those susceptible) were included in the empirical studies to which the model was fit. All persons in the S class can be infected at rate \( \lambda(t) \) (the force of infection) and move into the E class. They then progress from the E class into the I class (symptomatic) at a rate inversely proportional to the incubation period (1/\( \mu \)). We are thus assuming that when a susceptible (S) person becomes infected, disease will later develop and that all first infections are symptomatic. Persons then recover at a rate inversely proportional to duration of illness (1/\( \mu_i \)), at which point they are shedding asymmetrically (A). Infection then ends at a rate inversely proportional to duration of shedding (1/\( \rho \)), after which the person is assumed to have cleared the infection and is recovered from symptoms and that the person’s immune system protects from further disease (R).

<table>
<thead>
<tr>
<th>Study</th>
<th>No. challenged</th>
<th>No. (%) infected</th>
<th>No. (%) AGE</th>
<th>No. challenged</th>
<th>No. (%) infected</th>
<th>No. (%) AGE</th>
<th>No. challenged</th>
<th>No. (%) infected</th>
<th>No. (%) AGE</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolin 1971 (10)</td>
<td>12</td>
<td>9 (75)</td>
<td>SM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wyatt 1974 (11)†</td>
<td>23</td>
<td>16 (70)</td>
<td>NV, MC, HI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parrino 1977 (8)†</td>
<td>12</td>
<td>6 (50)</td>
<td>NV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson 1990 (17)†</td>
<td>42</td>
<td>31 (74)</td>
<td>NV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham 1994 (12)</td>
<td>50</td>
<td>34 (68)</td>
<td>NV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindesmith 2003 (18)</td>
<td>77</td>
<td>34 (44)</td>
<td>NV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindesmith 2005 (19)</td>
<td>15</td>
<td>9 (60)</td>
<td>7 (47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmar 2008 (20)</td>
<td>21</td>
<td>16 (76)</td>
<td>11 (52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leon 2011 (21)‡</td>
<td>15</td>
<td>7 (47)</td>
<td>5 (33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmar 2011 (14)‡</td>
<td>41</td>
<td>34 (83)</td>
<td>29 (71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seitz 2011 (22)</td>
<td>13</td>
<td>10 (77)</td>
<td>10 (77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French 2012 (23)</td>
<td>40</td>
<td>17 (42)</td>
<td>12 (30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*AGE, acute gastroenteritis; SM, Snow Mountain virus; NV, Norwalk virus; MC, Montgomery County virus; HI, Hawaii virus; GII.4, genogroup 2 type 4.
†Only includes initial challenge, not subsequent re-challenge.
‡Only includes placebo or control group.
Consistent with the understanding of NoV host response, in our model, NoV-specific immunity is not lifelong and we allow 2 pathways out of R class. First, persons can become asymptomatically infected by cycling back into the A class at the same force of infection to which S persons are subjected \( \lambda(t) \). As such, R class represents a type of immunity in which persons are subject to infection but not disease—they can become asymptotically infected and shed virus in stool specimens, but symptoms of AGE do not develop. Persons in R class can also lose their immunity to disease through the waning process, whereby they become fully susceptible again at a rate of \( 1/\theta \). \( \theta \) is a fitted parameter (described below). Births and deaths are assumed to be equal and occur at a constant rate throughout the year. Static model inputs are detailed in Table 2.

In this baseline model (Figure 1), we assume that only symptomatic (I) persons contribute to transmission, so \( \lambda(t) \) is a function of the number of susceptible persons, the age-specific contact rate \( \beta_i \), the prevalence of infection \( I(t) \), and the probability of transmission, given contact (see online Technical Appendix). We allow for children <5 years old to have a different, presumably higher, level of infectiousness \( q_i \) than older children and adults \( q_2 \) (Table 3).

### Model Scenarios

Our first model incorporated several simplifications (e.g., that the entire population is genetically susceptible) for which considerable uncertainty exists (e.g., that immunity to 1 strain of NoV protects against other strains). Therefore, we set up several scenarios to explore the effects on duration of immunity estimates of pre- and post-infection immunity.

### Table 2. Fixed input parameters for each model scenario for duration of immunity to norovirus gastroenteritis*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life expectancy, y</td>
<td>NA</td>
<td>76</td>
<td>76</td>
<td>76</td>
<td>76</td>
<td>76</td>
<td>76</td>
<td>CDC FastStats (24)</td>
</tr>
<tr>
<td>Duration of incubation, d</td>
<td>( \mu_s )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Atmar et al., 2008 (20)</td>
</tr>
<tr>
<td>Duration of symptoms, d</td>
<td>( \mu_a )</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>Atmar et al., 2008 (20)</td>
</tr>
<tr>
<td>Duration of asymptomatic infection, d</td>
<td>( \rho )</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>Rockx et al., 2002 (25)</td>
</tr>
<tr>
<td>Relative infectiousness during incubation period</td>
<td>NA</td>
<td>0</td>
<td>0.05</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Sukhrie et al., 2010 (26)</td>
</tr>
<tr>
<td>Relative infectiousness during asymptomatic infection period</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>Sukhrie et al., 2010 (26)</td>
</tr>
<tr>
<td>Proportion of population genetically resistant</td>
<td>( \tau )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Lindesmith et al., 2003 (18)</td>
</tr>
<tr>
<td>Strains included</td>
<td></td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>GII.4 only</td>
<td>All</td>
<td>Rosenthal et al., 2011 (27)</td>
</tr>
<tr>
<td>Boosting of immunity by asymptomatic infection?</td>
<td>NA</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>GII.4 only</td>
<td>No</td>
</tr>
</tbody>
</table>

*NA, not applicable; CDC, Centers for Disease Control and Prevention. Model A, only symptomatic infectiousness; model B, presymptomatic and postsymptomatic infectiousness (low); model C, presymptomatic and postsymptomatic infectiousness (high); model D, innate genetic resistance; model E, genogroup 2 type 4 (GII.4); model F, no immune boosting by asymptomatic infection.
postsymptomatic infectiousness, genetic resistance within a portion of the population, and whether immunity to NoV is strain specific (Table 2).

**Model A: Symptomatic Individuals Infectious**

In model A, described in the previous section, only symptomatic individuals are infectious. This model provides the basis for the 5 following iterations.

**Model B: Presymptomatic and Postsymptomatic Infectiousness (Low)**

Presymptomatic persons (E) have been observed to transmit NoV (28), although how often this occurs is not known. Also, exposed, but not-yet-symptomatic, persons (E) are 5% as infectious as symptomatic persons (26). Because they incubate the virus for only 1 day (1/2 as long as the symptomatic phase), they are 2.5% as infectious as a symptomatic case-patient over the course of their incubation period. Persons may shed virus after resolution of symptoms and may also become infected and shed virus without exhibiting symptoms. Again, their importance in transmission has not been quantified. Sukhrie et al. have demonstrated that asymptomatic shedders can transmit the virus, but they do so at lower levels than symptomatic persons (26,29). In this scenario, asymptomatic (A) and presymptomatic (E) persons are 5% as infectious as symptomatic persons. Because the mean duration of shedding is 10 days, asymptomatic and presymptomatic persons have a cumulative infectiousness of 25% compared with symptomatic persons (Table 2).

**Model C: Presymptomatic and Postsymptomatic Infectiousness (High)**

This model has the same structure as model B. However, persons in the exposed (E) and asymptomatic (A) compartments are 25% as infectious as symptomatic persons.

**Model D: Innate Genetic Resistance**

In model D, we assume that 20% of the population is completely resistant to infection and disease (i.e., they have the nonsecretor phenotype), and therefore play no role in the transmission process (Figure 1) (18). They do, however, continue to make contact with other persons and are included in empirical incidence estimates, so the whole population is included in this model, even though 20% cannot become infected. This model includes a separate class of persons born completely resistant to infection and disease (i.e., they have complete genetic resistance (G)).

**Model E: Genogroup 2 Type 4 (GII.4)**

In Models A–D, we assume that all NoVs are antigenically indistinguishable, since the degree of strain specificity of NoV immunity is not well understood. Model E tests the sensitivity of that assumption by including only GII.4 infections, which have been the predominant circulating strain for the past decade. We multiplied incidence data by 0.72 (an estimate of the proportion of all NoV AGE caused by GII.4 viruses) (32) to represent only GII-4 cases and subsequently refitted the model. This model assumes that GII.4 viruses are antigenically distinct from non-GII.4 NoVs and that all GII.4 viruses are antigenically indistinguishable (33).

**Model F: No Immune Boosting by Asymptomatic Infection**

Persons do not move from the recovered (R) to asymptomatic (A) compartments. The only pathway out of the R class is through waning of immunity to become susceptible (S) again.

**Data and Model Fitting**

We fit the model to age-specific incidence from the Study of Infectious Intestinal Disease in England (5) and

---

Table 3. Duration of immunity, fitted parameter estimates, and log-likelihood and basic reproductive number for models of duration of immunity to norovirus gastroenteritis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Model A</th>
<th>Model B</th>
<th>Model C</th>
<th>Model D</th>
<th>Model E</th>
<th>Model F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of immunity, y</td>
<td>$\theta$</td>
<td>5.1 (3.9–6.5)</td>
<td>5.1 (4.0–6.7)</td>
<td>8.7 (6.8–11.3)</td>
<td>4.1 (3.2–5.1)</td>
<td>7.6 (5.6–8.0)</td>
<td>5.1 (3.9–6.6)</td>
</tr>
<tr>
<td>Probability of transmission per</td>
<td>$q_t$</td>
<td>0.25</td>
<td>0.18</td>
<td>0.37</td>
<td>0.35</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>infected contact, 0–4 y</td>
<td></td>
<td>(0.21–0.31)</td>
<td>(0.15–0.21)</td>
<td>(0.14–0.91)</td>
<td>(0.27–0.44)</td>
<td>(0.19–0.25)</td>
<td>(0.21–0.31)</td>
</tr>
<tr>
<td>Probability of transmission per</td>
<td>$q_s$</td>
<td>0.050</td>
<td>0.036</td>
<td>0.094</td>
<td>0.062</td>
<td>0.051</td>
<td>0.050</td>
</tr>
<tr>
<td>infected contact, ≥5 y</td>
<td></td>
<td>(0.042–0.055)</td>
<td>(0.032–0.039)</td>
<td>(0.078–0.114)</td>
<td>(0.057–0.066)</td>
<td>(0.047–0.056)</td>
<td>(0.046–0.054)</td>
</tr>
<tr>
<td>Negative log likelihood</td>
<td></td>
<td>615.497</td>
<td>613.905</td>
<td>663.052</td>
<td>616.597</td>
<td>611.509</td>
<td>615.375</td>
</tr>
<tr>
<td>Annual incidence, %†</td>
<td></td>
<td>5.2</td>
<td>5.3</td>
<td>5.5</td>
<td>5.1</td>
<td>3.8</td>
<td>5.2</td>
</tr>
<tr>
<td>Basic reproductive number (all ages)</td>
<td>$R_0$</td>
<td>1.79</td>
<td>1.64</td>
<td>7.16</td>
<td>1.88</td>
<td>1.73</td>
<td>1.79</td>
</tr>
<tr>
<td>Basic reproductive number (0–4 y)</td>
<td>$R_0$</td>
<td>4.33</td>
<td>3.98</td>
<td>15.22</td>
<td>4.84</td>
<td>3.98</td>
<td>4.33</td>
</tr>
</tbody>
</table>

*Model A, only symptomatic infectiousness; model B, presymptomatic and postsymptomatic infectiousness (low); model C, presymptomatic and postsymptomatic infectiousness (high); model D, innate genetic resistance; model E, genogroup 2 type 4 (GII.4); model F, no immune boosting by asymptomatic infection.

†Compared with an observed annual incidence 4.5% from Phillips et al. (5), except for model E, which should be compared with norovirus GII.4-specific incidence of 3.2%.
the size of the adult (defined as 15–44 years of age) population immune at endemic equilibrium by allowing the transmission probabilities \((q_\text{s})\) and duration of immunity \((\theta)\) to vary during the fitting process. Size of the immune population was estimated from a literature review of challenge studies (Table 1).

We calculated the log-likelihood of the data under each model by assuming Poisson distributions with mean equal to the number of model-predicted cases for symptomatic NoV incidence in each age group and number of immune persons in the adult age group (see online Technical Appendix). Both incidence and population immune were treated as count data, on the basis of the size of the study population in the study in England (5) and the cumulative number of subjects included in challenge studies. The best-fitting parameter set maximized the log-likelihood of the age-stratified time series for the given set of estimated and fixed parameters (34). We calculated 95% CIs for each parameter (in each model) and generated a likelihood profile by holding a given parameter constant at a series of values and refitting the model. The upper and lower values were found by using the likelihood ratio test to determine at which parameter value the model converged on a significantly worse fit.

Because seasonality is a defining characteristic of NoV infection, we added seasonal forcing variables to visually inspect whether US outbreak patterns as described by Yen et al. (35) could be captured. We allowed the transmission coefficient \((\beta_1)\) to vary by 6% over the course of the year. However, because including seasonality did not qualitatively change our estimate of the duration of immunity, we excluded it in favor of a more parsimonious model.

**Results**

All models provided a qualitatively good fit to the crude incidence data, ranging from 5.1% (models D and E) to 5.5% (model C) per year, compared with the observed 4.5% per year (Table 3; Figure 2; online Technical Appendix Figure). All models also captured the decreasing incidence by age; model B was best able to represent the overall incidence and the high incidence in children <5 years of age (21.4% observed; 19.3% fitted), and model B roughly captured the incidence in the groups \(\geq 45\) years of age. Model C provided a worse fit than models A, B, D, or F. Model E could not be readily compared because it is fitted to a different incidence case count. Although model B was not a significantly better fit than A, D, or F, it did have the smallest negative log-likelihood, so we used model B for subsequent results, unless stated otherwise.

The \(R_0\) (basic reproductive number) for all models ranged from 1.64 to 1.88, except in model C, which had an \(R_0\) of 7.16. \(R_0\) for children 0–4 years of age was 15.22, substantially higher than for persons \(>5\) years \((R_0 = 0.89)\) (model B, Table 3).

In model A, the duration of immunity to NoV was estimated at 5.1 years (95% CI 3.9–6.5; Table 3). The duration of immunity estimated in model B was essentially the same as in model A at 5.1 years (95% CI 4.0–7.6). When the infectiousness of asymptomatic persons was increased in model C, estimate of duration of immunity increased substantially, to 8.7 (95% CI 4.0–11.3). Duration of immunity estimated in model D, in which transmission was effectively restricted to 80% of the population, was 4.1 years (95% CI 3.2–5.1). In model E, which was essentially fitted to a lower incidence to reflect only GII.4 transmission, duration of immunity was estimated at 7.6 years (95% CI 5.6–8.0). Model F, which did not allow subclinical infection to boost immunity, resulted in a duration of immunity estimate of 5.1 years (95% CI 3.9–6.6). Note that the transmission parameters \((q_\text{s})\) fell into 3 relative patterns: lower (model B), middle (models A, E, and F), and high (models C and D). These differences in transmissibility partly explain why the duration of immunity estimates are not more divergent between models.

With mild seasonal forcing (6% seasonal variation in transmission probabilities), the model captures fluctuations in disease incidence similar to those reported from outbreaks in 30 US states during 2007–2010 (Figure 3). Seventy-three percent of cases were estimated to occur during October–March, compared with 73% observed in the United States during October–March from 2007 to 2010.

**Discussion**

The goal of this study was to gain a better estimate of the duration of immunity to NoV AGE, and our results
suggest that it is longer than was previously understood. We modeled a range of possible infection and immunity processes to capture the unknown aspects of the transmission process, and from these models, we estimate a mean duration of immunity ranging from ≈4 to 8 years. Variations in duration of immunity between models can be traced to inclusion of presymptomatic and postsymptomatic states in the model scenarios. The best-fit models converge on a much higher infectiousness and $R_0$ for young children (<5 years) than for older children and adults. This finding is consistent with observational studies that found contact with a symptomatic child to be the prime risk factor for NoV infection for children and adults (5,30) and suggests that young children have a key role in the transmission of NoV to all age groups. Children have relatively high rates of contact with both other children and adults, and because of their lower levels of hygiene, they are likely to be more infectious than adults, given contact (31).

In our study, the models produced strong quantitative fits to the empirical data on incidence and population immunity, as well as on seasonality. Results suggest that parameter estimates are not overly sensitive to structural uncertainties, such as the role of asymptomatic shedding in disease transmission, at least within the range of the fixed parameters we have considered. The possible exception is model C, which resulted in a much higher $R_0$ than the other models and previous estimates, suggesting that asymptomatic persons are unlikely to be as infectious as they were parameterized to be in this scenario. Because the exact structure of immunity in the general population is unknown, our study sought to elucidate that structure rather than identify exact values for the various parameters included in each model scenario.

Several caveats should be borne in mind when interpreting these results. First, perhaps most critical, our model assumes that immunity is to disease (e.g., symptoms) rather than to infection. As such, so-called immune persons are still subject to becoming infected but they do not show symptoms. Infection without symptoms is a common outcome of exposure, as shown by volunteer studies and point prevalence of asymptomatic infection detected in the general population, which can be as high as 30% (5,18,19). In effect, our model allows for boosting of immunity by cycling between the recovered (R) and infectious asymptomatic (A) compartments. However, our estimates of duration of immunity pertain to time spent in the immune state from time of most recent symptomatic infection. If a person repeatedly became asymptotically infected (moves from R to A class), that person would effectively be immune to disease for longer than a person without successive asymptomatic/subclinical infections. The duration of immunity estimates are therefore conservative with respect to total time a person is protected from disease.

Second, with this single-strain model, we assume that all NoVs are antigenically indistinguishable and that infection with 1 NoV provides protection against all others. This is not strictly true (11), but data are not available on cross-protection to a range of NoV strains circulating at a particular time. As an extreme simplification of this process, we modeled GII.4 viruses on the assumption that they comprise 72% of observed incidence and are an antigenically homotypic genotype, essentially acting as a separate virus. However, GII.4 viruses are antigenically distinct from other GII viruses, and every few years, new GII.4 strains emerge that escape acquired population immunity. Over the past 15 years, at least 2 immune escape variants of GII.4 have emerged (in 2002 and 2006) (33). Although our estimate of duration of immunity (>4 years) may be compromised by this assumption, that novel GII.4s emerge once every 4 years or so would still suggest a role for the duration of immunity on the scale of years. Immunity gained through exposure to the prevalent strain would persist past emergence of a new strain, even though such protection could be effectively useless against the new strain.

![Figure 3. Norovirus gastroenteritis outbreak patterns from 30 US states, January 2007–April 2010 (black bars) and predicted annualized monthly incidence for all age groups (red line). These results are for model B (which includes presymptomatic and postsymptomatic infectiousness) and, for this illustration, seasonal forcing (35).](image)
These findings could ultimately have implications for vaccine policy. Empirical studies strongly document that children have the highest incidence of disease. Our results suggest that young children play a dominant role in the transmission process. Therefore, vaccinating young children is likely to result in both the greatest direct and indirect benefits. This conclusion is at odds with the current direction of vaccine development, which is increasingly focused on demonstrating safety and efficacy in older age groups (36,37). Future modeling studies could explicitly examine the potential direct and indirect benefits of vaccinating different age groups. Moreover, severe disease disproportionately occurs among the elderly (despite their lower incidence of disease), but the elderly are difficult to successfully immunize for both programmatic and immunologic reasons. Therefore, future modeling studies should address the question of whether severe disease outcomes could best be prevented directly, by vaccinating the elderly, or indirectly, by vaccinating children (38–40). Our study provides estimates of the infectiousness of children <5 years of age and adults (with the former being much more infectious) on which to base such simulations.

Because these results suggest a longer duration of protection than previously estimated, they support the continued development of NoV vaccines. A short duration of protection (<1 year, for example) would be a major impediment for widespread use of a NoV vaccine because it would have to be given frequently, and the distribution would be expensive and logistically difficult (e.g., willingness for annual vaccination). However, if duration of immunity and possibly vaccine protection are indeed on the order of 5 years, as this study suggests, the cost-benefits and health gains per person vaccinated could be substantially greater than previously estimated (41).

Our findings represent a substantial departure from current estimates of the duration of immunity to NoV. As noted, our models make several potentially influential simplifying assumptions. However, these models, grounded in observational evidence on age-specific incidence, seasonality of disease, and levels of population immunity, may be more realistic than results of re-challenge studies, which have formed the basis of current estimates. Specifically, this analysis suggests that the large dose or type (GI.1) delivered to volunteers in the classic challenge studies was unrepresentative of natural exposure to common contemporary strains. Because a robust duration of protection is likely crucial for the success of vaccines, future trials could consider following-up at least a subset of participants for several years either for natural disease or by challenge, providing an empirical test of these modeling results.

Funding for this project was provided by the Rollins Practical Experience Program and the CDC Foundation. In addition, this work was supported by grant 1K1087724 from the National Institute for Allergy and Infectious Diseases, National Institutes of Health, and a grant from Emory University Global Health Institute (to J.S.L.).

Ms Simmons is an Oak Ridge Institute for Science and Education (ORISE) Fellow at the Centers for Disease Control and Prevention, Atlanta. Her research interests include modeling social networks and infectious disease transmission.

**References**

Duration of Immunity to Norovirus Gastroenteritis

Technical Appendix

This Technical Appendix describes the differential equations used in a deterministic dynamic transmission model with age structure that tracks the population with respect to norovirus infection and immunity status.

Model Equations

Our model is a system of ordinary differential equations, described as follows:

\[
\frac{dG_i}{dt} = B\tau - D\tau
\]
\[
\tau \text{ is 0 in all models except E, so there is no genetically resist class in all other models.}
\]

\[
\frac{dS_i}{dt} = B(1 - \tau) + 1/\theta R_i - (\lambda i(t) + D)Si
\]

\[
\frac{dE_i}{dt} = \lambda(t)S_i - (1/\mu_s + D)Ei
\]

\[
\frac{dIS_i}{dt} = 1/\mu_s E_i - (1/\mu_a + D)IS_i
\]

\[
\frac{dIA_i}{dt} = 1/\mu_a IS_i + \lambda i(t)R_i - (1/\rho + D)IAi
\]
\[
\text{The } \lambda i(t)R_i \text{ term is not included in model F}
\]

\[
\frac{dR_i}{dt} = 1/\rho IA_i - (\lambda i(t) + 1/\theta + D)R_i
\]
\[
\text{The } \lambda i(t)R_i \text{ term is not included in model F}
\]

where:

\[
i = \text{age group, defined as 0-4 years, 5-14 years, 15-45 years and 45+ years}
\]

\[
G_i = \text{individuals genetically resistant to NoV}
\]

\[
S_i = \text{individuals susceptible to NoV}
\]
E_i = individuals exposed to NoV
S_i = individuals infected with NoV and showing symptoms
A_i = individuals infected with NoV and not showing symptoms, but shedding virus in stool
R_i = individuals who have recovered from NoV and are no longer shedding virus
τ = proportion of individuals who are genetically resistant to NoV
B = number of births entering the system
D = number of deaths exiting the system
θ = the rate at which recovered individuals become susceptible again; the duration of immunity
λ(t) = the force of infection; the rate at which susceptible or recovered individuals become infected
μ_e = the rate at which exposed individuals become infected symptomatic
μ_a = the rate at which infected symptomatic individuals become infected asymptomatic
ρ = the rate at which infected asymptomatic individuals recover

Seasonality

Norovirus has such a seasonal pattern that it was initially described as ‘the winter vomiting disease.’ Incidence peaks in the late winter and early spring, and most outbreaks are seen during this time. We accounted seasonal variation in transmissibility in the model as follows:

β(t) = β_0 ((1 + β_1 × cos(2πt + ω))

where β_0 is the mean of the transmission parameter, β_1 is the amplitude of its seasonal fluctuation and ω is the phase angle in years (t). The mean transmission parameter (β_1) depends on age-specific mixing and contact patterns of the population.

Force of infection (per Susceptible)

Since we had several scenarios with varying compartments for infected individuals, the expression for force of infection varied by scenario. These expressions are described below:

Scenarios A, D, E & F:
\[ \lambda_i(t) = \sum_{j=1}^{4} q_{ij}(t) c_{ij}(t) I_i = \sum_{j=1}^{4} \beta_{ij}(t) I_i \]

Scenario B:

\[ \lambda_i(t) = \sum_{j=1}^{4} q_{ij}(t)(0.05E_i + I_i + 0.05A_i) = \sum_{j=1}^{4} \beta_{ij}(t)(0.05E_i + I_i + 0.05A_i) \]

Scenario C:

\[ \lambda_i(t) = \sum_{j=1}^{4} q_{ij}(t)(0.25E_i + I_i + 0.25A_i) = \sum_{j=1}^{4} \beta_{ij}(t)(0.25E_i + I_i + 0.25A_i) \]

where \( i \) represents the various age groups and 0.05 or 0.025 is the value used to account for the differences in infectiousness between symptomatic individuals and those in the exposed and asymptomatic compartments. \( c_{ij} \) represents the contact rates between age group \( i \) and \( j \), such that \( \beta_{ij}(t) = q_{ij}(t) \) and is the rate at which two individuals come into effective contact with each other per unit time. All other terms are defined elsewhere in the manuscript or appendix.

### Contact Rates

Contact rates were obtained from the European POLYMOD study (1) and adjusted to account for the age ranges in our model. We specifically used social contact rates from Great Britain. The rates used in the model are detailed in the Table below.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>0-4 years</th>
<th>5-14 years</th>
<th>15-44 years</th>
<th>45+ years</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4 years</td>
<td>7.13621e-7</td>
<td>1.81279e-7</td>
<td>1.67541e-7</td>
<td>7.03979e-8</td>
</tr>
<tr>
<td>5-14 years</td>
<td>1.81279e-7</td>
<td>1.1697e-6</td>
<td>2.2533e-7</td>
<td>9.5557e-8</td>
</tr>
<tr>
<td>15-44 years</td>
<td>1.67541e-7</td>
<td>2.2533e-7</td>
<td>3.51465e-7</td>
<td>1.65856e-7</td>
</tr>
<tr>
<td>45+ years</td>
<td>7.03979e-8</td>
<td>9.5557e-8</td>
<td>1.65856e-7</td>
<td>2.03e-7</td>
</tr>
</tbody>
</table>

### Fitting

The model fit was assessed by calculating the log-likelihood of the data under the assumption that the observations (monthly incidence in each age class \((i=1..4)\) and the proportion immune \((i=5)\) at time \( t (d_{i,t}) \)) follow a Poisson distribution with mean \( d_{i,t} \) as follows:
The best-fitting model was the one for which the set of estimated parameters \( \{ \theta, q_1, q_2 \} \) yielded the maximum log-likelihood. We fit the model after an initial burn-in period of 50 years to allow the model to reach equilibrium.

**Calculation of \( R_0 \)**

We follow the approach outlined in the paper by Diekmann et al [1] to calculate the basic reproduction number for each of the models detailed in this study. This method requires us to construct a Next Generation Matrix (NGM) for our model system (outlined in the 'Model Equations' section of the Appendix) and calculate the eigenvalues of this matrix. \( R_0 \) is then equal to the largest eigenvalue, or spectral radius, of the matrix. We go about constructing the NGM by first reducing the system to the infected states (i.e. the compartments labeled E, I, and A) so that we have a system with 12 state variables consisting of three infected states for each of the four age-groups. We then populate a Transmission matrix (T)—which includes all epidemiological ‘births’—and a Transition matrix (E)—which contains all flows between infected compartments.

If we first define a 'force of infection' matrix \( F \) as:

\[
F = \begin{bmatrix}
q_1 \cdot \beta_{11} & q_2 \cdot \beta_{12} & q_2 \cdot \beta_{13} & q_2 \cdot \beta_{14} \\
q_1 \cdot \beta_{21} & q_2 \cdot \beta_{22} & q_2 \cdot \beta_{23} & q_2 \cdot \beta_{24} \\
q_1 \cdot \beta_{31} & q_2 \cdot \beta_{32} & q_2 \cdot \beta_{33} & q_2 \cdot \beta_{34} \\
q_1 \cdot \beta_{41} & q_2 \cdot \beta_{42} & q_2 \cdot \beta_{43} & q_2 \cdot \beta_{44}
\end{bmatrix}
\]

Where the values \( \beta_{ij} \) are the elements of the age-dependent mixing matrix and the values \( q_1 \) and \( q_2 \) are the infectivity of those under and over 5 years, respectively, then the Transmission matrix \( T \) can be written as:

\[
T = \begin{bmatrix}
\text{inf}_1 \cdot F & \text{inf}_2 \cdot F & \text{inf}_3 \cdot F \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

where 0 indicates a 4 by 4 matrix of zeros, and \( \text{inf}_1, \text{inf}_2 \) and \( \text{inf}_3 \) are the infectivity levels of each of the individuals in each of the infected compartments E, I, and A respectively.
And if we define:

\[ \begin{bmatrix} \mu_S & 0 & 0 & 0 \\ 0 & \mu_S & 0 & 0 \\ 0 & 0 & \mu_S & 0 \\ 0 & 0 & 0 & \mu_S \end{bmatrix}, \begin{bmatrix} \mu_A & 0 & 0 & 0 \\ 0 & \mu_A & 0 & 0 \\ 0 & 0 & \mu_A & 0 \\ 0 & 0 & 0 & \mu_A \end{bmatrix} \text{ and } \rho = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & \rho \end{bmatrix} \]

Our Transition matrix E is given by:

\[ E = \begin{bmatrix} -\mu_S & 0 & 0 \\ \mu_S & -\mu_A & 0 \\ 0 & \mu_A & -\rho \end{bmatrix} \]

The NGM K, is formulated as:

\[ K = -T. E^{-1} \]

and the basic reproduction number is calculated as the largest eigenvalue of K.

**Model fits to all scenarios**

Technical Appendix Figure. Observed annual incidence (black) compared with output from the model (red bars) by age group, displayed by scenario.
Reference