The pathogenicity of norovirus is definitively established. However, norovirus is frequently detected in the stool of healthy individuals. To gain understanding of the apparent high prevalence of asymptomatic infection, we analyzed a dynamic transmission model of norovirus infection, disease, and immunity. We simulated norovirus epidemiology in low- and high-transmission settings by varying the basic reproduction number ($R_0$). We predicted annual disease incidence values in children aged 0–4 years of 25% with a low $R_0$ and 29% with a high $R_0$. However, the point prevalence of asymptomatic infection rose sharply from 3% to 48% from the low to high $R_0$ settings. Among older children and adults, the models projected that incidence of disease would rise from 6% to 16% from the low to high $R_0$ settings, whereas asymptomatic infection prevalence was lower in this age group. Asymptomatic prevalence of norovirus can change dramatically with small changes in $R_0$. The ratio of prevalence in cases to controls could be high in a developed country and close to or even less than 1 in a high-exposure setting, despite similar disease incidence. These findings highlight an important limitation of case-control studies for pathogens for which there is suboptimal diagnostic specificity.

Abbreviation: RT-qPCR, reverse transcription–quantitative polymerase chain reaction.

Noroviruses are well recognized as the most common cause of acute gastroenteritis across all age groups (1). They are now the most common cause of pediatric diarrhea in the United States (2). Volunteers who are fed the virus experience vomiting and diarrhea and shed the virus in copious amounts (3). These symptoms are accompanied by pathophysiological changes, including blunting of the villi in epithelium of the intestinal wall (4). Taken together, these observations clearly demonstrate the pathogenicity of norovirus.

However, features of norovirus infection and human immunity make it difficult to interpret the significance of the presence of virus in stool of people with disease. A wide variety of factors may explain why individuals without diarrhea shed enteric pathogens in their stool (5). The factors most germane to norovirus are as follows. First, some individuals have true asymptomatic infection; 15%–35% of serologically confirmed or stool-confirmed infections in studies of adult volunteers are not associated with gastroenteritis (3, 6–8). These asymptomatic infections may result from acquired immunity that is protective against disease but does not block infection. Second, shedding of virus in stool continues long after the resolution of symptoms, with around 25% of cases shedding virus at detectable levels 3 weeks after the onset of illness, which usually lasts for only 1–3 days (9). The quantity of shed virus decreases 3–4 days postinfection, making viral load (based on cycle threshold value from quantitative reverse transcription–polymerase chain reaction (RT-qPCR)) an indicator of disease-causing infection (10). However, there is not a clear cutoff of viral load associated with symptoms; moreover, true asymptomatically infected individuals appear to initially shed virus at levels similar to norovirus gastroenteritis cases (3). Finally, it is possible that ingested virus, likely from small inocula, can cause subclinical infection or may even transit the gut without replicating or causing infection while still being detectable with highly sensitive diagnostics.
For this combination of reasons, norovirus can frequently be detected in stool collected from healthy individuals. In fact, in a number of studies, particularly in low-income countries, norovirus has been nearly as prevalent, and sometimes more so, in controls than in cases (Figure 1). This includes the recent Global Enteric Multicenter Study, the largest systematic assessment for understanding the etiology of childhood diarrhea in developing countries (11). The Global Enteric Multicenter Study and other case-control studies use the odds ratio of a microbe being present in cases versus healthy controls (12–14), which essentially quantifies the magnitude of the association between microbe and disease, to calculate an attributable or etiological fraction (13, 15). Accordingly, these studies would conclude that norovirus is a minor pathogen or, sometimes, not a pathogen at all. Among the 7 Global Enteric Multicenter Study sites in low-income African and Asian countries, norovirus was found to be consistently responsible for a significant attributable fraction of diarrheal disease in only 1 site (Basse, The Gambia) (11).

How do we reconcile these findings with overwhelming evidence that norovirus is a key agent of gastroenteritis? (4) The body of norovirus case-control study results provides a clear example of the limitations of this study design for a pathogen that commonly causes reinfection and for which the immune state of the individual determines whether infection or disease develops upon exposure. Mathematical models of infectious disease transmission are mechanistic and explicitly track the immune state of the population, so models can facilitate insights into the dynamic relationship between host immune status and exposure to a pathogen. To gain understanding of the apparently high prevalence of asymptomatic infection, we analyzed a dynamic transmission model of norovirus infection, immunity, and disease.

METHODS

Full details of the model can be found elsewhere (16). Here, we briefly describe the key attributes pertinent to the question of apparent asymptomatic infection. The model assumes that children are born susceptible to infection (Figure 2). The first infection always results in symptomatic disease, after which individuals shed virus for a period of time before becoming immune to symptomatic disease. “Immune” individuals can still become infected asymptotically with the same force of infection as fully susceptible individuals. As such, immunity is presumed to be against disease, not infection. Thus, individuals can cycle through the immune and asymptptomatically infected states if they continue to be exposed to the virus. Eventually, immunity can wane, and individuals become fully susceptible to disease again. We assume that the duration of asymptomatic shedding is the same as the duration of postsymptomatic shedding. This is based on the finding of Leon et al. (17) that virus could be detected in the stool of asymptomatic volunteers (n = 8) for a median of 21 days (range, 3–32) after resolution of symptoms and for 19 days (range, 14–23) in asymptatically infected volunteers (n = 4). Contact rates were obtained from the European POLYMOD Study (1) and adapted to the age ranges in our model. We used social contact rates from Great Britain.

Fitted and fixed parameters are given in the Appendix Table 1. We also considered an iteration of the model in which immune individuals are protected against disease and partially protected (50%) against infection. However, this model was unable to capture the high levels of asymptomatic infection (Figure 1) so was not further analyzed. Using maximum likelihood, we fit the model to age-specific incidence from the Infectious Intestinal Disease Study in England (5) and the size of the adult (ages 15–44 years) population immune at endemic equilibrium by allowing the
transmission probabilities \( (q, s) \) and the duration of immunity \( (1/\theta) \) to vary during the fitting process. Separate transmission probabilities were fitted for transmission from those aged 0–4 years and those aged 5 years and older.

There are 2 main parameters for considering asymptomatic infection prevalence in this framework. The first is \( R_0 \), the number of secondary cases arising from the average primary case in a fully susceptible population, a measure of transmission that differs between populations. For \( R_0 \), we take 3 values to represent low-transmission settings (fitted to England’s age-specific incidence rates) and medium- and high-transmission settings (by doubling and tripling the value; \( R_0 = 1.64, 3.28 \), and 4.92 respectively). There are neither \( R_0 \) estimates nor incidence estimates for norovirus in developing countries, so we base this conservative assumption on \( R_0 \) differentials for poliovirus, another enteric virus once common in all settings; Fine and Carneiro (18) summarized estimates of \( R_0 \) from 2 to 4 in areas of “good hygiene” and from 8 to 14 in areas of “poor hygiene.” The second factor is the duration of detectable shedding in asymptomatic infection. Duration of shedding is important for understanding asymptomatic prevalence because, all else being equal, if shedding lasts for 10 days, the point prevalence in the population will be double what it would be if shedding lasted 5 days. Our predicted relationship is tautological; a longer duration of shedding results in a higher point prevalence of asymptomatic infection. In other words, prevalence is a function of incidence multiplied by duration (in this case, of shedding).

We also considered, to a limited extent, the effect of demography by reducing life expectancy by 10 and 20 years in the middle- and low-income scenario settings, respectively. We found very little difference in either symptomatic disease incidence or asymptomatic prevalence (<1% difference compared with the results presented below). However, a full assessment of the effect of realistic demography in a particular setting is beyond the scope of this paper. Such an analysis would need to include age-specific death rates, local contact patterns, and norovirus incidence data to fit the model in lower-income settings. The latter 2 data elements are lacking, thus precluding such an analysis at this time.

We define the point prevalence of asymptomatic shedding as \( A_i / N_i \), where \( A_i \) is the number of individuals in age group \( i \) shedding norovirus and not showing symptoms, and \( N_i \) is the population in that age group. Seasonal forcing was removed from the model (same \( R_0 \) all year); the model was run to endemic equilibrium.

RESULTS

We predicted annual disease incidence in children aged 0–4 years of 25% with a low \( R_0 \) and 33% with a medium \( R_0 \), a marginal increase, followed by a decrease to 29% with a high \( R_0 \) (Figure 2). However, the point prevalence of asymptomatic infection rose sharply from 2.8% to 21% to 48% from the low to high \( R_0 \) settings, assuming a duration of postsymptomatic shedding of 15 days. Among older children and adults, the models projected that incidence of disease would rise sharply from 6.1% to 14.9% to 15.6% from the low- to high-transmission settings, whereas asymptomatic infection prevalence was negligible (<1%) in low \( R_0 \) settings and rose to 3.8% in medium-transmission and 20% in high-transmission settings. These estimates are consistent with the range of asymptomatic prevalence detected in field studies (as in Figure 1), as well as the few studies reporting age-specific disease incidence (19). We present results in broad age categories (stratified at age 5 years) because results were similar (in terms of infection and disease) among finer older-age groups (5–14, 15–44, and 45 years and older).

Underlying these results is an epidemiologic scenario in which children have high levels of exposure, such that they have high levels of disease, but they also have high levels of immunity to disease and therefore frequently develop asymptomatic infection. We estimate that, at equilibrium, 62% of children in low \( R_0 \) settings have acquired immunity as a result of their high levels of exposure. However, immunity wanes at a rate of 1/5.1 years, so older groups are projected to have lower levels of immunity (between 27% and 48%). With increasing \( R_0 \), more individuals across the age range are exposed to norovirus. However, because children’s immune status, most of their additional exposure results in asymptomatic infection, whereas the majority of additional exposure in older individuals results in symptomatic disease. In the medium- to high-transmission settings, we project that more than 75% of children aged 0–4 years have acquired immunity because they are exposed to norovirus soon after birth. Figures 3 and 4 also illustrate that there is an upper limit of symptomatic disease incidence in children (projected to be ~33% annually). Asymptomatic prevalence, on the other hand, could approach 100% given frequent enough exposure.
Figure 4 also presents how the duration of infection (which here can also be interpreted as the duration of detectable shedding) is a strong driver of asymptomatic prevalence. In the low $R_0$ scenario, asymptomatic prevalence increases from less than 1% (at 5 days) to 18% (at 30 days); in the high $R_0$ scenario, asymptomatic prevalence increases from 4.8% (at 5 days) to 80% (at 30 days) among children aged 0–4 years (Figure 4A). Analogous patterns, at lower asymptomatic infection levels, are projected for the population aged 5 years and older (Figure 4B).

**DISCUSSION**

Our modeling framework, in which norovirus is a pathogen that confers partial immunity, predicts high prevalence of norovirus among gastroenteritis-free individuals in high-transmission settings. These results have practical implications concerning the interpretation of case-control studies of norovirus and other infections that frequently result in asymptomatic excretion. Prevalence of asymptomatic norovirus infection can change dramatically with fairly small changes in $R_0$. Therefore, prevalence in healthy individuals (controls) may increase, while the case incidence stays relatively constant. For this reason, the odds ratio (or pathogenicity index or attributable fraction) for norovirus, as calculated in a case-control study of the etiology of diarrheal disease, may be high in a developed country and close to unity, or even less than 1, in a high-exposure setting, despite similar disease incidence. An extreme example of this discrepancy is the 21%:4% case-to-control prevalence ratio in medically attended gastroenteritis in children in the United States (2) compared with 22%:31% in children in Botswana (20). It should be noted, however, that low prevalence has not always been detected among asymptomatic individuals in studies conducted in high-income settings. In the Infectious Intestinal Disease Study, 50% of cases and 29% of controls in England tested positive for norovirus (21). The Infectious Intestinal Disease Study was one of the first to use RT-qPCR, which is now the diagnostic standard for norovirus. Cases tended to have higher viral loads than controls, though the distributions were overlapping (10), so probabilities of the likelihood that norovirus was “disease-causing” at a certain level were estimated (19) as opposed to using a clear diagnostic cutoff.

Although imperfect diagnostic accuracy is an issue for all assays, the problem can be more acute for RT-qPCR. Indeed, RT-qPCR is able to detect norovirus at a load as low as $10^5$ copies per gram of stool, so its exquisite analytical sensitivity compromises diagnostic specificity. As such, low viral loads consisting of nonreplicating virus particles can be detected by RT-qPCR. Rotavirus diagnostics provide a clear example of the utility of an appropriately sensitive and specific diagnostic assay. The limit of detection of enzyme immunoassays used for the diagnosis of rotavirus correlates with levels at which rotavirus tends to be disease causing (22). However, when more sensitive RT-qPCR assays are used (typically for research rather than clinical diagnosis), the association between detection of virus and disease symptoms becomes much weaker (22, 23). Future studies should be designed with sufficiently large sample sizes to determine if viral load measured by RT-qPCR as its proxy can be used to identify disease-causing norovirus infections; these studies are specifically needed in developing countries and in older age groups where data are limited.

The Global Enteric Multicenter Study chose to use a conventional multiplex reverse transcription–polymerase chain reaction assay for the detection of several enteric RNA viruses, including norovirus, which was a decision based on
performance, robustness, cost-effectiveness, and expert judgment (24). Using this assay, researchers found that norovirus was a significant contributor to diarrheal disease in only 1 of the 7 sites (Basse, The Gambia) (11). We have attempted to illustrate the reasons that a case-control study conducted with a diagnostic assay of imperfect sensitivity and specificity may fundamentally underestimate the role of norovirus as a cause of diarrhea, especially in high-transmission settings. Another possibility is that norovirus is not a major cause of severe diarrheal disease (and is restricted to mild illness), and that norovirus diarrhea cases rarely present for medical care in developing countries. In this scenario, norovirus infection is purely coincidental to infection with another causative pathogen. However, we find it highly unlikely that norovirus can be the primary cause of medically attended diarrhea in the United States (2) and not be a cause of severe disease among children in the developing world. Our model presents a more likely explanation, which is that common reinfection in developing countries complicates the relationship between norovirus detection and disease.

Enteric pathogens do not typically confer lifelong sterilizing immunity and, therefore, reinfection is common. Determining the incidence of asymptomatic infection is challenging in epidemiologic studies, which are typically designed with disease endpoints, but for some pathogens, infection may be much more common than disease. A recent seroepidemiologic study, for example, found that the incidence of asymptomatic infection of Campylobacter is orders of magnitude greater than reported disease rates (25). Indeed, using Campylobacter as an example, Swart et al. (26) illustrated that, because of the interplay between the waning and boosting of immunity, reducing the force of infection does not necessarily lead to reduction in disease incidence. Our results for norovirus are in agreement, as shown by the predicted drop in incidence in children aged 0–4 years in medium-transmission compared with high-transmission settings. In addition, we find that this effect can be age dependent. Reducing the force of infection may not have much of an effect on incidence in children, but it would in adults. For a pathogen that does not confer sterilizing immunity, exposure can develop into either asymptomatic infection or disease; the outcome depends on the immune state of the individual.

As a final methodological point, the prevalence of a pathogen among cases is a rather different measure than the prevalence in healthy controls. The former is a function of the incidence of a given pathogen relative to all other causes of the disease. The latter is a reflection of the asymptomatic prevalence, which can be a combination of postasymptomatic shedding, true asymptomatic shedding, and even detection of nonreplicating virus transiting the gut. In the absence of an assay that can discriminate between true disease and background shedding, comparing prevalence in cases with prevalence in controls may have little value. Birth cohort studies with prospective follow-up will be needed to answer questions about the burden of disease from norovirus. Such studies could elucidate how primary and subsequent infections lead to the development of immunity, and whether immunity protects against infection or only the manifestation of disease symptoms.

ACKNOWLEDGMENTS

Author affiliations: Division of Viral Diseases, National Center for Immunization and Respiratory Disease, Centers for Disease Control and Prevention, Atlanta, Georgia (Ben Lopman, Kirsten Simmons, Manoj Gambhir, Jan Vinjé, Umesh Parashar); Rollins School of Public Health, Emory University, Atlanta, Georgia (Ben Lopman, Kirsten Simmons); and MRC Center for Outbreaks Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom (Manoj Gambhir).

No specific funding was used for this project. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Conflict of interest: none declared.

REFERENCES

Appendix Table 1. Fixed and Fitted Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life expectancy</td>
<td>Years</td>
<td>L</td>
<td>76</td>
</tr>
<tr>
<td>Duration of incubation</td>
<td>Days</td>
<td>1/(\alpha_s)</td>
<td>1</td>
</tr>
<tr>
<td>Duration of symptoms</td>
<td>Days</td>
<td>1/(\mu_a)</td>
<td>2</td>
</tr>
<tr>
<td>Duration of asymptomatic infection</td>
<td>Days</td>
<td>1/(\rho)</td>
<td>15</td>
</tr>
<tr>
<td>Relative infectiousness during incubation period</td>
<td>NA</td>
<td>(\alpha_E)</td>
<td>0.05</td>
</tr>
<tr>
<td>Relative infectiousness during asymptomatic infection period</td>
<td>NA</td>
<td>(\alpha_A)</td>
<td>0.05</td>
</tr>
<tr>
<td>Fitted parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of immunity</td>
<td>Years</td>
<td>1/(\theta)</td>
<td>5.1</td>
</tr>
<tr>
<td>Probability of transmission per infectious contact in those aged 0–4 years</td>
<td>NA</td>
<td>(q_1)</td>
<td>0.18</td>
</tr>
<tr>
<td>Probability of transmission per infectious contact in those aged 5 years and older</td>
<td>NA</td>
<td>(q_2)</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.