Taking Account of Asymptomatic Infections in Modeling the Transmission Potential of the COVID-19 Outbreak on the Diamond Princess Cruise Ship

April 18, 2020

Li-Shan Huang, Ph.D., Professor of Statistics, National Tsing Hua University, Hsinchu, Taiwan
Li Li, M.A., Lead Financial Analyst, AT&T, Bedminster, NJ
Lucia Dunn, Ph.D., Professor Emerita of Economics, Ohio State University, Columbus, Ohio
Mai He, M.D., Ph.D., Associate Professor of Pathology and Immunology, Washington University School of Medicine in St. Louis

Abstract

We take the case of the Diamond Princess (DP) cruise ship as an experimental model for studying the transmission potential of COVID-19 in a closed environment. We investigate the changes in $R_0$ for COVID-19 on the DP from January 21 to February 19, 2020 with a chain-binomial model at different times under two scenarios: no quarantine assuming a random mixing condition, and quarantine of passengers in cabins — passengers may get infected either by an infectious case in a shared cabin or by asymptomatic crew who continued to work. Our modeling approach takes account of the asymptomatic ratio of crew members during the quarantine of passengers, which has not been explored in the literature. Assuming an asymptomatic ratio 0.505 and the proportion of infections in cabins 0.2, $R_0$ at the beginning of the epidemic was 3.27 (95% CI (3.02, 3.54)) and 3.78 (95% CI (3.49, 4.09)) respectively for serial intervals of 5 and 6 days, and increased for passengers in contact with asymptomatic crew during quarantine. We find evidence to support a CDC report that “a high proportion of asymptomatic infections could partially explain the high attack rate among cruise ship passengers and crew.” This emphasizes the importance of improved reporting and quarantine of asymptomatic cases, and raises questions on quarantine procedures in closed environments such as military vessels, cruise ships, dormitories, prisons, and other enclosed living complexes with high population densities.
Introduction

This COVID-19 outbreak has developed into an international public health emergency [1]. The reproductive number ($R_0$) of COVID-19 is a key piece of information for understanding an epidemic. Current intervention methods focus on quarantine methods with either mitigation or suppression strategies aimed at reducing the reproduction number $R_0$ and flattening the curve [2]. Asymptomatic infectious cases are less likely to seek medical care or to be tested and quarantined, contributing to the infectious potential of a respiratory virus [3, 4]. Clinical findings have suggested that the viral load in asymptomatic patients is similar to that in symptomatic patients [5]. Evidence suggests that these asymptomatic patients can infect others before they manifest any symptoms [6, 7, 8]. In an early study of cases in Wuhan [9], 200 individuals out of 240 (83%) reported no exposure to an individual with respiratory symptoms, which suggests presymptomatic/asymptomatic infection is common [10]. The DP data [11, 12, 13] allow us to further examine the impact of asymptomatic cases in a closed environment. A CDC report [14] states that “a high proportion of asymptomatic infections could partially explain the high attack rate among cruise ship passengers and crew.”

The DP, with 3,711 people on board as of February 5, 2020, was found to have an outbreak of COVID-19 from one traceable passenger from Hong Kong. This passenger became symptomatic on January 23 and disembarked on January 25 in Hong Kong. On February 1, six days after leaving the ship, he tested positive for SARS-CoV-2 at a Hong Kong hospital. Japanese authorities were informed about this test result. On February 4, the authorities announced positive test results for SARS-CoV-2 for another ten people on board. The ship was quarantined by the Japanese Ministry of Health, Labour and Welfare for what was expected to be a 14-day period, off the Port of Yokohama [11]. Initially, passengers were quarantined in their cabins while the crew continued to work. Only symptomatic cases and close contacts were tested for COVID-19 and PCR-confirmed positive passengers were removed and isolated in Japanese hospitals. As reported [12], attempts were made to test all passengers including asymptomatic cases starting on February 11, and as of February 20, 619 cases had been confirmed (16.7% of the population on board), including 82 crew and 537 passengers. Overall, 712 (19.2%) of the crew and passengers tested positive [14]. Since 50.5% of the COVID-19 cases on the DP were asymptomatic [12], this situation is very close to a real-life scenario that would be found in urban areas.

The $R_0$ of COVID-19 on DP has been estimated previously in [15]; they identified the $R_0$ as 14.8 initially and then declining to a stable 1.78 after the quarantine and removal interventions. They also suggest that the $R_0$ of COVID-19 and contact rate are dependent on population density. That research does not take account of asymptomatic cases. Other researchers using data on the DP up
to February 16 have estimated the median $R_0$ as 2.28 [16]. They found $R_0$ remained high despite quarantine measures, while concluding that estimating $R_0$ was challenging due to the difficulty in identifying the exact number of infected cases.

We investigated the changes in $R_0$ for COVID-19 on the DP from January 21 to February 19 with a chain-binomial model at different times under two scenarios: no quarantine assuming a random mixing condition, and quarantine of passengers in cabins — passengers may get infected either by an infectious case in a shared cabin or by asymptomatic crew who continued to work. Our modeling approach takes account of the asymptomatic ratio of crew members in the quarantine of passengers, which has not been explored in the literature.

**Materials and Methods**

**Data Sources**

The Diamond Princess data from January 21 to February 19 were taken from the Japan National Institute of Infectious Disease website [11, 12, 13]. We set January 21 as day 1, since January 20 was the start date (day 0) of the cruise. February 19 (day 30) was the date that most passengers were allowed to leave the ship. For those dates that $Y$, the number of new COVID-19 cases, was not reported, linear interpolation was used. As an example, there were 67 new cases on February 15, but no data were reported on February 14. After linear interpolation, $Y$ on a daily basis became 33 and 34 for February 14 and 15 respectively. Based on the documented onset dates [11], there were 34 cases with onset dates before February 6, and we further adjusted the number of confirmed cases on February 3, 6, and 7, from 10, 10, and 41 cases to 17, 17, and 27 cases respectively. We chose serial intervals $\tau$ of 5 and 6 days as these are factors of 30 and are close to 7.5 days (95% CI 5.3 to 19) in [9] and 4 days in [17, 18]. Then daily data were aggregated into 5- and 6-day intervals.

**Statistical Analysis**

The chain-binomial model originally proposed in [19] belongs to the broader class of stochastic discrete-time SIR models [20]. The model assumes that an epidemic is formed from a succession of generations of infectious individuals from a binomial distribution. For the DP data, the initial population size is $N_{t=0} = 3711$, where time $t$ is the duration measured in units of the serial interval. To model the dynamics on the ship for the case $\tau = 6$, we make the following assumptions.

(a) From January 21 to 26 ($t = 1$, the first serial interval), infection contacts happened at random following the random mixing assumption. Let $I_t$ be the number of persons infected
at time \( t \). Then \( I_{t=1} \) is a binomial random variable \( B(N_{t=0}, p_1) \) with binomial transmission probability \( p_1 = 1 - \exp(-\beta \times I_{t=0}/N_{t=0}) \), where \( \beta \) is the transmission rate and \( I_{t=0} = 1 \) (the first case who disembarked on January 25). As in the SIR model, the probability that a subject escapes infectious contact is assumed to be \( \exp(-\beta \times I_{t=0}/N_{t=0}) \).

(b) From January 27 to February 1 \((t = 2)\), infection contacts again followed the random mixing assumption. Hence \( I_{t=2} \) is a binomial random variable \( B(N_{t=0}, p_2) \) with \( p_2 = 1 - \exp(-\beta \times I_{t=0}/N_{t=0}) \), and the number of persons at risk of infection is \( N_{t=1} = 3711 - I_{t=1} \).

(c) For the period February 2 to 7 \((t = 3)\), \( I_{t=3} \) is a binomial random variable \( B(N_{t=2}, p_3) \) with \( N_{t=2} = N_{t=1} - I_{t=2} \) and \( p_3 = 1 - \exp(-\beta \times (I_{t=1} + I_{t=2})/N_{t=0}) \). As the quarantine started on February 5 and confirmed cases were removed, the number of persons infected, removed, and at risk of infection at the end of the \( t=3 \) period were \( I_{t=3} \), \( I_{t=1} + I_{t=2} \), and \( N_{t=3} = N_{t=2} - I_{t=3} \) respectively.

(d) For \( t = 4 \) and \( t = 5 \) during the quarantine of passengers, \( N_t = N_{t-1} - I_t \), and we further make the following assumptions. (i) Of all infected cases, 86.8% were passengers and 13.2% crew [11]. We use these proportions to calculate the number of infected persons in each group, \( I_{p_t} \) and \( I_{c_t} \), respectively, and \( I_t = I_{p_t} + I_{c_t} \). This assumption is imposed since there is no public data available for the time course of \( I_{p_t} \) and \( I_{c_t} \). (ii) Crew members continued to work unless showing symptoms; hence the binomial transmission probability of \( I_{c_t} \) remained the same as \( 1 - \exp(-\beta \times I_{t=1}/N_{t-1}) \), \( t = 4 \) and 5. In other words, crew were randomly mixing in the population on board. (iii) Passengers stayed in cabins most of the time. Assume that among infected passengers \( I_{p_t} \), the proportion of infections that occurred in cabins is \( r_p \), and that the average occupancy per cabin is 2. For those \( I_{p_t-1} \) cases, \( t = 4, 5 \), the binomial transmission probability to infect \( I_{p_t} \times r_p \) passengers is \( 1 - \exp(-\beta/2) \). Based on [11], \( r_p = 0.2 \) (= 23/115), while we assume \( r_p = 0.2 \) and 0.3 when \( t = 6 \). For this assumption, \( r_p \leq I_{p_{t-1}}/I_{p_t} \), and thus \( r_p \) cannot be made arbitrarily large. (iv) The other \((1 - r_p)\) proportion of infected passengers’ cases was possibly due to asymptomatic crew members who continued to perform service [11], and their binomial transmission probability is assumed to be \( p_t = 1 - \exp(-\beta \times aratio \times I_{c_{t-1}}/C_{t-1}) \), where \( aratio \) is the asymptomatic ratio and \( C_{t-1} \) is the number of crew members on board at time \( t - 1 \). That is to say, these passengers were randomly mixing in the crew population with possible infectious contact with asymptomatic crew. In our calculations, \( aratio = 0.4, 0.465 [14], 0.505 [12], 0.6, and 0.7. \)

Assumptions (a)–(c) correspond to no quarantine assuming a random mixing condition, and assumption (d) to quarantine of passengers in cabins in which passengers may either get infected (d)(iii) by an infectious case in a shared cabin or (d)(iv) by asymptomatic crew who continued to work. The maximum likelihood (ML) approach was used to estimate \( \beta \). We modified some of the
functions in [20] for the chain-binomial model and the ML step was carried out using the R-package bbmle [21]. For \( t = 4, 5 \), \( R_0 \) is the number of persons (passengers or crew) at risk (at time \( t \)) times either (d)(iv) \( 1 - \exp(-\beta \times \text{aratio}/C_{t-1}) \) for passengers potentially infected by asymptomatic crew, or (d)(ii) \( 1 - \exp(-\beta/N_{t-1}) \) for crew.

The calculation for the case \( \tau = 5 \) is analogous: period January 21 to 25 follows (a); period January 26 to 30 follows (b); period January 31 to February 4 follows (c); and periods February 5 to 9, February 10 to 14, and February 15 to 19 follow (d), and \( r_p = 0.2 \), which is the maximum value given the constraint \( r_p \leq I_{p-t-1}/I_{p-t} \).

**Results**

For the Diamond Princess COVID-19 outbreak, Table 1 gives the estimates of \( \beta \) and their 95% confidence intervals. Since \( \beta \) is the basic reproductive number \( R_0 \) at the beginning of the epidemic (\( t = 1, 2, 3 \) when \( \tau = 5 \), and \( t = 1, 2 \) when \( \tau = 6 \)), we observe from Table 1 that the estimated \( R_0 \) for the initial period is greater than 3 in every one of the scenarios that we considered.

\( R_0 \) as a function of \( t \) is illustrated in Figure 1 for \( \tau = 5 \) and \( \tau = 6 \) days, in the case where \( r_p = 0.2 \) and \( \text{aratio} = 0.505 \). Detailed results are as follows.

Based on the chain-binomial model, the values of \( R_0 \) as a function of time \( t = 3, 4, 5 \) when \( \tau = 6 \) are as follows. With \( r_p = 0.2 \), \( \text{aratio} = 0.505 \), and estimated \( \beta = 3.78 \) (Table 1), \( R_0 \) at \( t = 3 \) is approximately the same as those of \( t = 1, 2 \). \( R_0 \) for passengers in (d)(iv) is 4.73 (95% CI (4.37, 5.12)) and 4.39 (95%CI (4.06, 4.75)) at \( t = 4, 5 \) respectively, and 1.06 (95%CI (0.98, 1.15)) and 1.05 (95%CI (0.97, 1.14)) for crew in (d)(ii) respectively. This shows that \( R_0 \) for some passengers increased from 3.78 to 4.73 from \( t = 3 \) to \( t = 4 \) if they were in contact with asymptomatic crew members. From \( t = 4 \) to \( t = 5 \), \( R_0 \) for those passengers decreased slightly to 4.39, mostly due to removal of a larger number of infected passengers. The \( R_0 \) for crew at \( t = 4, 5 \) is small and close to 1, since most infected passengers were removed and crew were exposed to fewer cases. When \( \tau = 5 \), \( r_p = 0.2 \), \( \text{aratio} = 0.505 \), and estimated \( \beta = 3.27 \), the \( R_0 \) for passengers in (d)(iv) is 4.18 (95% CI (3.86, 4.52)), 4.08 (95%CI (3.77, 4.42)), and 3.74 (95% CI (3.45, 4.04)) at \( t = 4, 5 \), and 6 respectively, and for crew in (d)(ii) is 0.92 (95%CI (0.85, 1.00)), 0.92 (95%CI (0.85, 0.99)), and 0.91 (95% CI (0.84, 0.98)) respectively.

Other than those infections among passengers sharing the same cabin, when \( \tau = 6 \) days, the combined \( R_0 \) for passengers and crew is 2.90 (95%CI (2.67, 3.13)) and 2.73 (95%CI (2.52, 2.95)) respectively for \( t = 4 \) and 5, decreasing from the initial \( R_0 = 3.78 \), illustrating the effects of quarantine. Similarly, for \( \tau = 5 \) days, the combined \( R_0 \) for passengers and crew is 2.55 (95%CI
(2.36, 2.76)), 2.50 (95%CI (2.31, 2.71)), and 2.32 (95%CI (2.15, 2.52)) respectively for $t = 4$, 5 and 6.

Figure 2 shows 100 stochastic simulations of the chain binomial model based on $N = 3,700$ and estimated $\beta = 3.78$, assuming no quarantine, infected cases removed, $\tau = 6$, and extrapolation to 90 days, with the observed epidemic (red line). It indicates that the quarantine on the DP did prevent a more serious outbreak. If there was no quarantine, the cumulative number of cases at the end of 30 days has a mean 855 (SD = 439), and median 791 (IQR = 533), while the observed DP data of 621 cases is at the 35th percentile. Among 99 of 100 simulations, the entire population is infected at the end of 54 days.

**Discussion**

Due to the closed environment and close contact among people on board, cruise ships can be vulnerable to outbreak of infectious diseases [14]. Our estimates of $R_0$ for the initial period (Table 1) are all greater than 3, consistent with most estimates of $R_0$ reported earlier [22, 23], showing that the COVID-19 virus is highly transmissible.

Our results show that with a serial interval of 6 days, $R_0$ is similar for $t = 1–3$, yet $R_0$ for some passengers is slightly higher for $t = 4$, while $R_0$ finally decreases for $t = 5$. This demonstrates the possibility that as long as asymptomatic positive cases were not removed, $R_0$ remains high under quarantine conditions. We find evidence to support a CDC report that “a high proportion of asymptomatic infections could partially explain the high attack rate among cruise ship passengers and crew.”

The effects of pre/asymptomatic population on the spread of COVID-19 during quarantine has not yet been studied extensively. Clinical observations and lab tests have confirmed the existence of a pre/asymptomatic population infecting others [6, 7, 8]. It is not easy to give precise estimates of the size of this population, yet the DP outbreak provides very useful real world data for this. 46.5% of passengers and crew members on the DP were asymptomatic at the time of testing [14], and about 17.9% of the infected individuals never demonstrated any symptoms [13].

Some research has suggested that the pre/asymptomatic population, “silent carriers,” are the main driving force behind this pandemic. A group has estimated that the proportion of undocumented infections in China — including those who experience mild, limited or no symptoms and go undiagnosed— could be as high as 86% prior to January 23, 2020. They estimated the transmission rate of undocumented infections as 55% of the rate for documented infections, and yet that undocumented infections contributed to 79% of documented cases [23].
Another group found that the total contribution from the pre/asymptomatic population is more than that of symptomatic patients [25].

The Italian town Vò Euganeo of 3,300 people conducted blanket testing on all residents, with initial testing showing 3% positives. All positives were quarantined which included pre/asymptomatic cases; on the second round of testing, the number of positive results dropped significantly to 0.3%. The report of this intervention considers that quarantine of asymptomatic positives contributed to control of the infection [26]. In contrast, a number of other cruise ships and some military vessels, which were not able to perform immediate ship-wide screening, suffered rapid spread of COVID-19 similar to what occurred on the DP. [27, 28].

The limitations of this modeling study are as follows. First, due to inadequate data on the time course of infection cases among crew and passengers, assumption (d)(i) assumes a constant proportion, which may vary with time in practice. Second, the values of the parameter $r_p$ assumed in the present study may not be sufficiently large. Third, the investigation on asymptomatic COVID-19 cases is ongoing and it is unknown whether the values of the asymptomatic ratio used in this paper are close to the truth. Finally, the assumptions (a)-(d) under chain-binomial models may not be sufficient to capture the complexity of the COVID-19 epidemics. Fuller data reporting is important for researchers to develop statistical methodology to help combat this pandemic.

On the DP, crew members continued to perform service unless they showed symptoms. This provides a parallel to people doing “essential work” in society and thus exempt from shelter-in-place rules. These groups contain “silent carriers” which include the pre- and asymptomatic infected population and may be a major force in the COVID-19 pandemic. We suggest, with due caution, that it may be necessary to develop stronger protocols to address these cases, especially in confined spaces, in order to successfully control the spread of the virus.

References


Table 1: Estimates of β and their 95% confidence intervals for the Diamond Princess COVID-19 outbreak data.

<table>
<thead>
<tr>
<th>asymptomatic ratio</th>
<th>β = R₀ under initial conditions; estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>τ = 6; rₚ = 0.2</td>
</tr>
<tr>
<td>40%</td>
<td>3.94 (3.64, 4.27)</td>
</tr>
<tr>
<td>46.50%</td>
<td>3.84 (3.54, 4.16)</td>
</tr>
<tr>
<td>50.50%</td>
<td>3.78 (3.49, 4.09)</td>
</tr>
<tr>
<td>60%</td>
<td>3.64 (3.36, 3.94)</td>
</tr>
</tbody>
</table>

Figure 1: Time-dependent effective reproduction number R₀ (solid lines) of COVID-19 on board the Diamond Princess ship January 21 (day 1) to February 19 (day 30) and their 95% confidence intervals, assuming τ = 5 and 6 days, rₚ = 0.2 and aratio = 0.505. Red: R₀ for passengers in contact with asymptomatic crew members. Blue: R₀ for crew members.
Figure 2: 100 stochastic simulations of the chain binomial model based on $N = 3700$ and estimated $\beta = 3.78$, assuming no quarantine, infected cases removed, and extrapolation to 90 days, $\tau = 6$, with the observed epidemic (red line).