Research article

Estimating relative generation times and reproduction numbers of Omicron BA.1 and BA.2 with respect to Delta variant in Denmark

Kimihito Ito[1],*, Chayada Piantham[2] and Hiroshi Nishiura[3]

[1] International Institute for Zoonosis Control, Hokkaido University, Sapporo, Japan
[2] Graduate School of Infectious Diseases, Hokkaido University, Sapporo, Japan
[3] School of Public Health, Kyoto University, Kyoto, Japan

* Correspondence: Email: itok@czc.hokudai.ac.jp; Tel: +81-11-706-9504; Fax: +81-11-706-9491.

Abstract: The Omicron variant spreads fastest as ever among the severe acute respiratory syndrome coronaviruses 2 (SARS-CoV-2) we had so far. The BA.1 and BA.2 sublineages of Omicron are circulating worldwide and it is urgent to evaluate the transmission advantages of these sublineages. Using a mathematical model describing trajectories of variant frequencies that assumes a constant ratio in mean generation times and a constant ratio in effective reproduction numbers among variants, trajectories of variant frequencies in Denmark from November 22, 2021 to February 26, 2022 were analyzed. We found that the mean generation time of Omicron BA.1 is 0.44–0.46 times that of Delta and the effective reproduction number of Omicron BA.1 is 1.88–2.19 times larger than Delta under the epidemiological conditions at the time. We also found that the mean generation time of Omicron BA.2 is 0.76–0.80 times that of BA.1 and the effective reproduction number of Omicron BA.2 is 1.25–1.27 times larger than Omicron BA.1. These estimates on the ratio of mean generation times and the ratio of effective reproduction numbers have epidemiologically important implications. The contact tracing for Omicron BA.2 infections must be done more quickly than that for BA.1 to stop further infections by quarantine. In the Danish population, the control measures against Omicron BA.2 need to reduce 20–21% of additional contacts compared to that against BA.1.

Keywords: SARS-CoV-2; Omicron-BA.1; Omicron-BA.2; relative generation time; relative reproduction number

1. Introduction

The Omicron variant spreads fastest as ever among the severe acute respiratory syndrome coronaviruses 2 (SARS-CoV-2) we had so far. The variant was designated as a variant of concern by the World Health Organization (WHO) on November 26, 2021 [1]. As of February 7, 2022, infections by Omicron were reported to the WHO from official sources of 159 countries [2]. The rapid replacement of Delta by Omicron followed by a steep rise in SARS-CoV-2 infections has been observed after the introduction of Omicron in many countries, indicating that Omicron has considerably higher transmission advantages than Delta in the vaccinated population [3]. Furthermore, the BA.2 sublineage of Omicron is replacing the BA.1 sublineage in Denmark [4]. Although attenuated disease severity of Omicron was reported [5,6], there is an urgent need to evaluate the transmissibility of these Omicron sublineages.

Transmissibility of a new variant should be characterized by two factors associated with the transmission of the virus. The first factor is the effective reproduction number (R_t), which measures how many secondary cases are generated by a single primary case at time t. The second factor is the generation time distribution $f(\tau)$, which describes how secondary infections are distributed as a function of time τ since the infection of primary case [7].

Selby [8] has estimated the generation time of Omicron using case counts in England in 2021. In a logistic regression assuming that the generation time of Omicron is the same of that of Delta, the log odds of Omicron and Delta frequencies should be on a straight-line relationship. However, the case counts in England showed a bend in the log odds of observations that deviated from its theoretical straight line. Using this bend, Selby estimated the ratio of the mean generation time of Omicron to the mean generation time of Delta to be 0.46 (95% confidence intervals (CI): 0.38–0.61). Comparing transmission pairs of infections with S gene target failure (SGTF) viruses with those of non-SGTF viruses, Backer et al. [9] showed that Omicron BA.1 has shorter serial intervals that non-SGFT viruses such as Delta.

Leung et al. [10,11] developed a mathematical model describing trajectories of variant frequencies, assuming that ratios among variants' effective reproduction numbers were constant. This method, however, needs daily numbers of infections of each variant. In our previous paper, we have proposed an approximated form of the model which is free from using daily numbers of infections [12]. Using the proposed model in [12], we have analyzed nucleotide sequences of SARS-CoV-2 collected in Denmark [13]. This model, however, assumed that generation times of variants followed the same probability distribution and that effective reproduction numbers of Omicron BA.1 and BA.2 were the same. In this paper, we propose a new model in which the generation times of variants may be different from each other. Analyzing latest variant count data observed in Denmark using the new model, we compare generation times and effective reproduction numbers of Omicron BA.1 and BA.2 with those of Delta.

2. Materials and methods

2.1. Model of variant frequencies

Consider a situation where viruses of variant a are circulating in the population at the beginning of the target period of analysis and variants A_1 and A_2 are newly introduced to the same population at

calendar times t_{A_1} and t_{A_2}, respectively. Let $q_a(t)$, $q_{A_1}(t)$, and $q_{A_2}(t)$ be frequencies of variants a, A_1, and A_2 in the viral population at calendar time t, respectively. Since there is no infection of A_1 before its introduction, $q_{A_1}(t)=0$ for a calendar time $t < t_{A_1}$, and the same is true for A_2 for $t < t_{A_2}$.

The dynamics on the frequencies of infections by a, A_1, and A_2 is determined by the length of generation times and the transmissibility, i.e., effective reproduction numbers of the three variants. We assume that the generation times (GT) of infections by A_1 and A_2 are c_1 and c_2 times longer than that of a, respectively. We call the value of c_i the relative generation time of A_i with respect to (w.r.t.) a. Let $f_a(\tau), f_{A_1}(\tau)$, and $f_{A_2}(\tau)$ be probability density functions of generation time distributions of infections by a, A_1, and A_2, respectively. The assumption on generation times can be described by the following equation:

$$c_i f_{A_i}(c_i \tau) = f_a(\tau) \tag{1}$$

for any generation time $\tau \geq 0$. To deal with the calendar time system, we truncate the generation time distributions at $\tau = 1$ and $\tau = l$ and discretize it as follows:

$$g_\chi = \begin{cases} 0, & \text{if } j = 0 \\ \int_{j-1}^{j} f_\chi(\tau)d\tau \Big/ \int_0^l f_\chi(\tau)d\tau, & \text{if } 1 \leq j \leq l \\ 0, & \text{if } j > l \end{cases} \tag{2}$$

for each variant χ in a, A_1, and A_2.

Let $I(t)$ be the total number of new infections by either a, A_1, or A_2 at calendar time t, and let $R_a(t)$, $R_{A_1}(t)$, and $R_{A_2}(t)$ be effective reproduction numbers of variants a, A_1, and A_2 at calendar time t, respectively. From the definition of the instantaneous reproduction number [7], $R_a(t)$, $R_{A_1}(t)$, and $R_{A_2}(t)$ can be written as follows:

$$R_a(t) = \frac{q_a(t)I(t)}{\sum_{j=1}^{l} g_a(j)q_a(t-j)I(t-j)}, \tag{3}$$

$$R_{A_i}(t) = \frac{q_{A_i}(t)I(t)}{\sum_{j=1}^{l} g_{A_i}(j)q_{A_i}(t-j)I(t-j)} \tag{4}$$

We assume that a patient infected by A_1 and A_2 generates respectively k_1 and k_2 times as many secondary transmissions as those of a patient infected by a regardless of time t. Using effective reproduction numbers of a, A_1, and A_2, this assumption can be described by the following equation:

$$R_{A_i}(t) = k_i R_a(t) \tag{5}$$

for each calendar time $t \geq t_{A_i}$. The value of the constant k_i is called relative reproduction number of A_i w.r.t. a.

Under the two epidemiological assumptions described by Equations (1) and (5), the frequency of variant A_i in the viral population at time t is now modeled as

$$q_{A_i}(t) = \frac{q_{A_i}(t)I(t)}{q_a(t)I(t) + \sum_{x=1}^{2} q_{A_x}(t)I(t)}$$

$$= \frac{R_{A_i}(t)\sum_{j=1}^{l} g_{A_i}(j)q_{A_i}(t-j)I(t-j)}{R_a(t)\sum_{j=1}^{l} g_a(j)q_a(t-j)I(t-j) + \sum_{x=1}^{2} R_{A_x}(t)\sum_{j=1}^{l} g_{A_x}(j)q_{A_x}(t-j)I(t-j)}$$

$$= \frac{k_i\sum_{j=1}^{l} g_{A_i}(j)q_{A_i}(t-j)I(t-j)}{\sum_{j=1}^{l} g_a(j)q_a(t-j)I(t-j) + \sum_{x=1}^{2} k_x\sum_{j=1}^{l} g_{A_x}(j)q_{A_x}(t-j)I(t-j)}. \tag{6}$$

Equation (6) needs to use $I(t)$, which are difficult to be quantified accurately in real time. To allow statistical estimation of c_i and k_i without knowing $I(t)$, we consider an assumption that the number of new infections does not greatly vary within a single generation of transmission from time $t-1$ to $t-l$, i.e.

$$I(t-1) \approx I(t-2) \approx \cdots \approx I(t-l), \tag{7}$$

for each calendar time t. Then we obtain our alternative formula:

$$q_{A_i}(t) = \frac{k_i\sum_{j=1}^{l} g_{A_i}(j)q_{A_i}(t-j)}{\sum_{j=1}^{l} g_a(j)q_a(t-j) + \sum_{x=1}^{2} k_x\sum_{j=1}^{l} g_{A_x}(j)q_{A_x}(t-j)}. \tag{8}$$

Note that Equation (7) is used just for approximating Equation (6) with Equation (8) for each calendar time t. We do not assume that $I(t)$ is constant during the entire period of the analysis, although a chain of the assumptions leads such a consequence. Equation (6) and Equation (8) will be compared using sequence data from Denmark.

2.2. Model of observation counts

Let $N_a(t)$, $N_{A_1}(t)$, and $N_{A_2}(t)$ be the number of variants a, A_1, and A_2 observed at calendar time t. We assume that variants a, A_1, and A_2 are sampled following a multinomial distribution with probabilities $q_a(t)$, $q_{A_1}(t)$, and $q_{A_2}(t)$ in $N_a(t) + N_{A_1}(t) + N_{A_2}(t)$ trials. Then parameters c_i and k_i and the initial variant frequencies of $q_{A_1}(t_{A_1})$ and $q_{A_2}(t_{A_2})$ can be estimated by maximizing the likelihood function of the multinomial distributions, which is described as follows:

$$L(c_1, c_2, k_1, k_2, q_{A_1}(t_{A_1}), q_{A_2}(t_{A_2}) ; N_a(t_{1:m}), N_{A_1}(t_{1:m}), N_{A_2}(t_{1:m}), I(t_{1:m}))$$

$$= \prod_{y=1}^{m} \frac{\left(N_a(t_y) + N_{A_1}(t_y) + N_{A_2}(t_y)\right)!}{N_a(t_y)! N_{A_1}(t_y)! N_{A_2}(t_y)!} q_a(t_y)^{N_a(t_y)} q_{A_1}(t_y)^{N_{A_1}(t_y)} q_{A_2}(t_y)^{N_{A_2}(t_y)}. \tag{9}$$

Note that Equation (8) does not use $I(t)$ even if they are given.

2.3. Observation of variants in Denmark

As of February 28, 2022, the earliest Danish Omicron sequences registered in the GISAID

database [14] were sampled on November 22, 2021. We downloaded a total of 150,164 nucleotide sequences of SARS-CoV-2 viruses sampled in Denmark during November 22, 2021 to February 22, 2022 from the GISAID database on February 28, 2022. Of these, 47,698 were labeled as PANGO lineages corresponding to Delta [15] and 102,222 were labeled as lineages corresponding to Omicron. Of sequences corresponding to Omicron, 36,118 were labeled as BA.1, 66,076 were BA.2, seven were BA.3, and 21 were B.1.1.529. The other 244 sequences were labeled as lineages corresponding to neither Delta nor Omicron. The 272 sequences that were not labeled as Delta, BA.1 or BA.2 were ignored in the subsequent analyses. We consider Delta as the baseline variant a and Omicron BA.1 and BA.2 as A_1 and A_2 in our model. We set t_{A_1} and t_{A_2} to November 22, 2021 and December 5, 2021, which are the sample collection dates of the earliest Omicron BA.1 and BA.2, respectively. For each date t from November 22, 2021 to February 22, 2022, the number of sequences labeled as Delta collected at that day was used as $N_a(t)$, that of Omicron BA.1 as $N_{A_1}(t)$ and that of Omicron BA.2 as $N_{A_2}(t)$. Dates of sample collection and PANGO lineages assigned to sequences are provided in Supplementary Table 1.

2.4. Numbers of confirmed cases in Denmark

Cumulative numbers of confirmed cases in Denmark were retrieved from Johns Hopkins Coronavirus Resource Center [16] on May 10, 2022. Daily changes in the cumulative numbers were recorded as raw daily confirmed cases $I_{raw}(t)$. For each date t during the period of the analysis, the 7-day rolling average of $I_{raw}(t)$ centered at t was recorded as average daily confirmed cases $I_{ave}(t)$. Some of the models used either $I_{raw}(t)$ or $I_{ave}(t)$ as the daily confirmed cases $I(t)$ in Equation (6).

2.5. Generation times of the Delta variant

Hart et al. reported that the distribution of the generation time of the Delta variant had a mean of 4.7 days with a standard deviation of 3.3 days by analyzing household transmission data from UK Health Security Agency [17]. Based on the mean and the standard deviation, the generation times of infections by Delta were assumed to follow the gamma distribution with a shape parameter of 2.03 and a scale parameter of 2.32. We used the probability density function of this gamma distribution as $f_a(\tau)$. The l was chosen to be 16, where is the 99th percentile of the generation time distribution of Delta.

2.6. Maximum likelihood estimation of parameters

We estimated the parameters of our models from the dataset using three assumptions on the generation time: The Three-GT, Two-GT, and One-GT assumptions. The Three-GT models assume that the GT distribution of Omicron BA.1 and BA.2 may be different from Delta in a manner described by Equation (1) and that the reproduction number of Omicron BA.1 and BA.2 may be different in a manner described by Equation (5). More precisely, c_1 may be different from c_2 and k_1 may be different from k_2. The Two-GT models assume that BA.1 and BA.2 share an identical generation time distribution. In other words, c_1 must be equal to c_2 and k_1 may be different from k_2. The One-GT models assume that the GT distributions of Omicron BA.1 and BA.2 infections are the same as that of Delta, i.e. $c_1 = c_2 = 1$, and that the reproduction number of Omicron may be different in a manner described by Equation (5). Note that $c_1 = c_2$ in the Three-GT assumption implies the Two-GT assumption, and

that $c_1 = c_2 = 1$ in the Two-GT assumption and Three-GT assumption implies the One-GT assumption. The maximum likelihood estimation and the 95% CI calculation were done using daily count data from November 22, 2021 to February 22, 2022. The Three-GT-raw-I model, Two-GT-raw-I model, and One-GT-raw-I model use $I_{raw}(t)$ as $I(t)$ in Equation (6) for the calculation of $q_\chi(t)$, under Three-GT, Two-GT, and One-GT assumptions, respectively. The Three-GT-ave-I model, Two-GT-ave-I model, and One-GT-ave-I model use $I_{ave}(t)$ as $I(t)$ in Equation (6), under Three-GT, Two-GT, and One-GT assumptions, respectively. The Three-GT-no-I model, Two-GT-no-I model, and One-GT-no-I model use Equation (8), which calculates $q_\chi(t)$ without $I(t)$, under Three-GT, Two-GT, and One-GT assumptions, respectively. The 95% CIs of parameters are derived from the profile likelihood method [18]. We used the augmented Lagrangian algorithm implemented in the NLopt module of the Julia language [19] for the maximum likelihood estimation and the calculation of 95% CIs of parameters. Models were compared using the Akaike information criterion (AIC) [20].

3. Results

The relative generation time of Omicron with respect to (w.r.t.) Delta, c_i, and the relative reproduction number k_i, were estimated from the daily count of Delta and Omicron BA.1 and BA.2 in Denmark using raw daily case information (Table 1), using 7-day rolling average case information (Table 2), and using no daily case information (Table 3). In all three Tables, the Three-GT assumption estimated smaller k_1 and k_2 compared to those of the Two-GT and One-GT assumptions.

Table 1. Results of parameter estimation with raw daily case information.

Parameter	Three-GT-raw-I model	Two-GT-raw-I model	One-GT-raw-I model
c_1 (95%CI)	0.58 (0.56–0.59)	0.59 (0.57–0.61)	1.00 (1.00–1.00)
c_2 (95%CI)	0.50 (0.48–0.51)	$= c_1$	$= c_1$
k_1 (95%CI)	2.12 (2.08–2.13)	2.13 (2.09–2.17)	2.66 (2.65–2.67)
k_2 (95%CI)	2.83 (2.76–2.86)	2.93 (2.87–3.01)	4.30 (4.27–4.34)
$q_{A_1}(t_{A_1})$ (95%CI)	$25.0\ (23.7\text{–}26.5) \times 10^{-4}$	$25.8\ (22.4\text{–}29.6) \times 10^{-4}$	$76.9\ (73.1\text{–}79.3) \times 10^{-4}$
$q_{A_2}(t_{A_2})$ (95%CI)	$6.6\ (5.6\text{–}7.8) \times 10^{-4}$	$13.6\ (12.7\text{–}14.6) \times 10^{-4}$	$15.9\ (14.5\text{–}17.3) \times 10^{-4}$

Table 2. Results of parameter estimation with 7-day rolling average daily case information.

Parameter	Three-GT-ave-I model	Two-GT-ave-I model	One-GT-ave -I model
c_1 (95%CI)	0.44 (0.43–0.46)	0.48 (0.46–0.51)	1.00 (1.00–1.00)
c_2 (95%CI)	0.36 (0.34–0.38)	$= c_1$	$= c_1$
k_1 (95%CI)	1.88 (1.85–1.91)	1.92 (1.88–1.97)	2.63 (2.61–2.65)
k_2 (95%CI)	2.35 (2.29–2.41)	2.51 (2.44–2.61)	4.24 (4.20–4.29)
$q_{A_1}(t_{A_1})$ (95%CI)	$16.3\ (15.2\text{–}17.9) \times 10^{-4}$	$18.8\ (16.4\text{–}21.3) \times 10^{-4}$	$76.4\ (71.5\text{–}83.4) \times 10^{-4}$
$q_{A_2}(t_{A_2})$ (95%CI)	$2.9\ (2.5\text{–}3.2) \times 10^{-4}$	$8.7\ (8.0\text{–}9.5) \times 10^{-4}$	$10.3\ (9.4\text{–}11.3) \times 10^{-4}$

Table 4 shows AIC of the nine models tested in this study. For each observation setting in raw-I, ave-I, and no-I, the Three-GT assumption has better AIC than the Two-GT assumption, and the Two-GT assumption has better AIC than the One-GT assumption. The values of AIC suggested that The Three-GT-no-I model, which is based on the Three-GT assumption and does not use $I(t)$, is the best

model to represent the sequence data in Denmark, followed by the Three-GT-ave-I model. The AIC of Three-GT-raw-I model was larger than those of Two-GT-no-I and Two-GT-ave-I models, indicating that the noise in raw case numbers made the likelihood of the model lower. From values of c_1, c_2, k_1, and k_2 in the Three-GT-no-I and Three-GT-ave-I models, it is suggested that the generation times of Omicron BA.2 is 0.76–0.80 times that of BA.1 and that the effective reproduction number of Omicron BA.2 is 1.25–1.27 times larger than that of Omicron BA.1 under the same epidemiological condition.

Table 3. Results of parameter estimation using no daily case information.

Parameter	Three-GT-no-I model	Two-GT-no-I model	One-GT-no-I model
c_1 (95%CI)	0.46 (0.45–0.47)	0.49 (0.46–0.51)	1.00 (1.00–1.00)
c_2 (95%CI)	0.35 (0.34–0.36)	$= c_1$	$= c_1$
k_1 (95%CI)	2.19 (2.18–2.21)	2.23 (2.18–2.27)	2.84 (2.81–2.86)
k_2 (95%CI)	2.79 (2.76–2.83)	2.95 (2.85–3.04)	4.70 (4.64–4.74)
$q_{A_1}(t_{A_1})$ (95%CI)	9.2 (8.6–10.2) $\times 10^{-4}$	10.9 (9.3–12.8) $\times 10^{-4}$	62.5 (58.5–68.6) $\times 10^{-4}$
$q_{A_2}(t_{A_2})$ (95%CI)	2.4 (2.1–2.9) $\times 10^{-4}$	9.1 (8.4–9.5) $\times 10^{-4}$	10.8 (9.9–11.9) $\times 10^{-4}$

Table 4. Akaike information criterion (AIC) of assumed models.

Model	Number of parameters	Maximum log likelihood	AIC
Three-GT-no-I	9	−650.8	1319.6
Three-GT-ave-I	9	−668.4	1354.7
Two-GT-no-I	8	−685.3	1386.6
Two-GT-ave-I	8	−699.5	1415.0
Three-GT-raw-I	9	−746.4	1510.7
Two-GT-raw-I	8	−759.9	1535.8
One-GT-no-I	7	−885.3	1784.6
One-GT-ave-I	7	−903.9	1821.8
One-GT-raw-I	7	−913.3	1840.6

Figure 1 shows observed frequencies and estimated frequencies of Delta, Omicron BA.1, and Omicron BA.2 estimated using the Three-GT-no-I model. For most dates in the analyzed period, the trajectories of variant frequencies estimated by the model were within 95% confidence intervals of population variant frequencies estimated by binomial tests using observed variant counts. This indicates that Equation (8) represents the time evolution of variant frequencies in the population quite accurately.
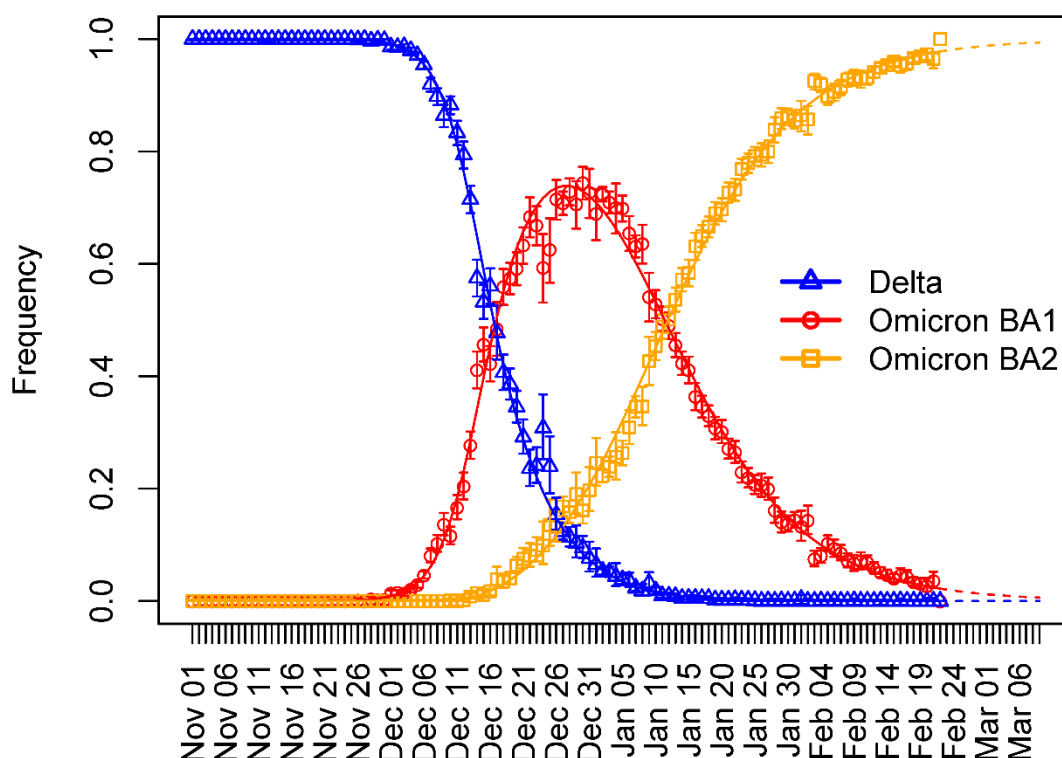
Figure 1. Frequencies of Delta, Omicron BA.1 and Omicron BA.2 from November 22, 2021 to March 9, 2022 estimated using the Three-GT-no-I model. Triangles, circles, squares represent the frequencies of Delta, Omicron BA.1, and BA.2, respectively. Solid lines indicate the maximum likelihood estimates of frequencies of Delta (blue), Omicron BA.1 (red), and Omicron BA.2 (orange). Dashed lines from February 23, 2022 to March 9, 2022 indicate predicted frequencies of Delta (blue), Omicron BA.1 (red), and Omicron BA.2 (orange). Error bars show the upper bound and lower bound of 95% confidence intervals of population frequencies of Delta (blue), Omicron BA.1 (red), and Omicron BA.2 (orange) estimated by binomial tests using observed counts. Error bars longer than 0.5 (the last data point) were not drawn.

Figure 2 shows the population average of the relative generation time and relative reproduction number of SARS-CoV-2 infections w.r.t Delta estimated from variant observations in Denmark using the Three-GT-no-I model. The replacement of Delta by Omicron BA.1 speeded up around December 7, 2021. From then, the population average of the generation times started decreasing and the population average of the relative reproduction numbers started increasing due to the Delta–Omicron BA.1 replacement. The decrease in the GT was saturated because Delta was being replaced by Omicron BA.1 and BA.2. However, population averages of relative reproduction numbers continued to increase due to the Omicron BA.1–Omicron BA.2 replacement.
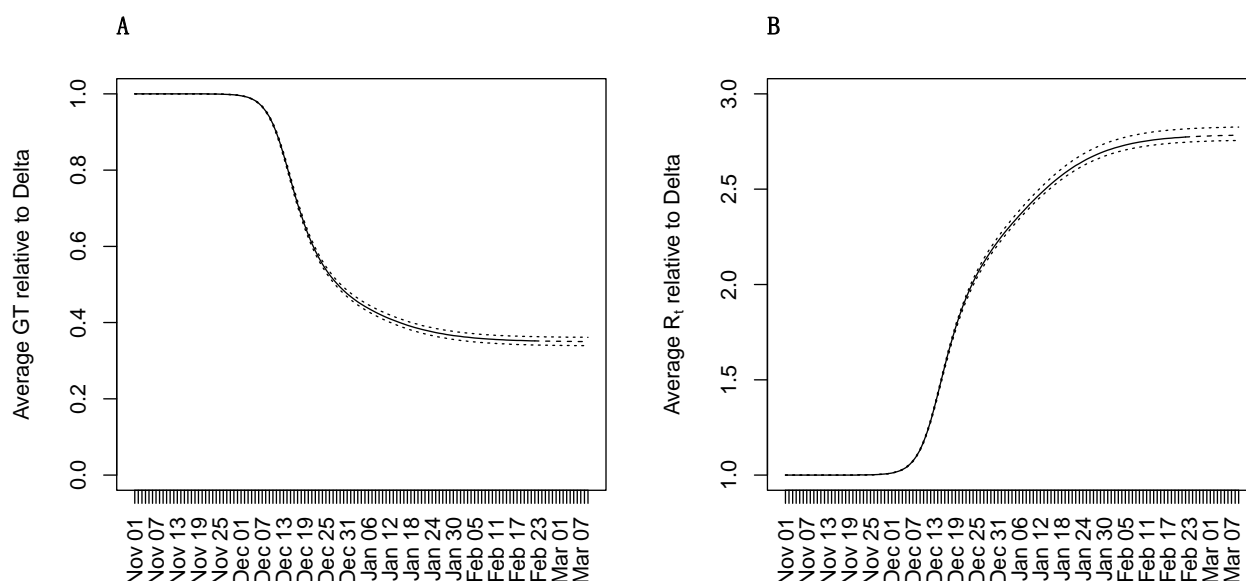
Figure 2. Population average of the relative generation time (A) and relative reproduction number (B) with respect to the Delta variant from November 22, 2021 to March 9, 2022 estimated by the Three-GT-no-I model. In each panel, the solid line represents maximum likelihood estimates until February 22, 2022 and the dashed line represents predicted values from February 23 to March 9, 2022. Dotted lines indicate the upper and lower bounds of the 95% confidence intervals of estimated and predicted values.

4. Discussion

We have developed a mathematical model describing the time evolution of variant frequencies using both the ratio of generation times and the ratio of effective reproduction number among variants. Analyzing variant counts of SARS-CoV-2 sampled in Denmark from November 22, 2021 to February 22, 2022 using the model, we have shown that the selective advantage of Omicron BA.1 and BA.2 over Delta was decomposed into advantage in the speed of transmission and advantage in the number of transmissions. We found that the mean generation time of Omicron BA.1 is 0.44–0.46 times that of Delta and the effective reproduction number of Omicron BA.1 is 1.88–2.19 times larger than that of Delta under the epidemiological conditions at the time. We also found that the generation times of Omicron BA.2 is 0.76–0.80 times that of BA.1 and that the effective reproduction number of Omicron BA.2 is 1.25–1.27 times larger than that of Omicron BA.1.

The estimation of these parameters has important implications to the control of current infections by Omicron BA.2. First, Omicron BA.2 generates 1.25–1.27 times more secondary transmissions than Omicron BA.1 in a partially immuned population in Denmark. Suppose BA.1 and BA.2 have effective reproduction numbers of $R_{BA.1}(t)$ and $R_{BA.2}(t)$ at time t, respectively. Control measures against Omicron BA.2 need to reduce 20%–21% of secondary infections to reduce $R_{BA.2}(t)$ to $R_{BA.1}(t)$. Second, generation times of Omicron BA.2 is 0.76–0.80 times that of Omicron BA.1. The contact tracing for Omicron BA.2 infections must be done more quickly than that for BA.1 to stop further infections by quarantine. Although the mean generation time of Omicron BA.2 is shorter than that of Omicron BA.1,

it is still necessary to investigate the infectious period of both variants to shorten the quarantine period for Omicron BA.2 infections.

In our previous paper, we have estimated the relative reproduction number of Omicron w.r.t. Delta to be 3.19 (95%CI: 2.82–3.61) using variant frequencies observed in Denmark from November 1, 2021 to December 9, 2021 [13]. The discrepancy in the relative reproduction numbers from this study is attributed to the following reason. Ito et al. 2021 assumed the same generation time for Omicron and Delta. In fact, the estimate using the One-GT-no-I model estimated relative reproduction numbers of 2.84 (2.81–2.86) for BA.1 and 4.70 (4.64–4.74) for the BA.2 w.r.t Delta (Table 3). Both BA.1 and BA.2 were circulating in Denmark in December 2021 [4], and the relative reproduction number estimated by our previous study is consistent with results of this study.

Table 4 suggested that the Three-GT-no-I model, which does not use $I(t)$, best represents observed data among the nine models tested in this study. It is known that the estimation of instantaneous reproduction number is sensitive to the noise in $I(t)$ [21]. In fact, the likelihood of Three-GT-raw-I model was lower than those of Two-GT-no-I and Two-GT-ave-I models. Use of the approximated version in Equation (8) instead of Equation (6) may have contributed to the elimination of the observation noise of $I(t)$. Use of the 7-day rolling average in $I(t)$ has also contributed to the noise reduction. Although the AIC supported the Three-GT-no-I model, the estimated values in Table 3 may be biased to some extent due to the approximation in Equation (7). On the other hand, the estimated values in Table 2 may also be biased due to the averaging operation for $I(t)$. It is difficult to judge whether estimates using the Three-GT-no-I model is more accurate than those using the Three-GT-ave-I model or not. One advantage to use the Three-GT-no-I model over the Three-GT-ave-I model is that one can predict the time course of variant replacement in future [22]. However, further studies are required to evaluate the bias introduced by the approximation in Equation (7).

We assume that the observational noise follows the multinomial distribution in Equation (9). This implies that the multinomial sampling from the population is assumed to be the only source of the noise. However, the noise may come from other possible sources, such as the regional difference in variant distributions. The narrow confidence intervals of estimated parameters may be attributed to the simple error model, and the true uncertainty may be larger than our confidence intervals. Such extra noise can be evaluated by using Dirichlet-multinomial distribution [22].

In our analyses, we set the introduction date of a variant to the date when the variant was first observed in the sequence dataset. We could estimate 95% confidence intervals of the introduction dates of variants; however, we have not done this because the estimation of introduction dates is not the purpose of this study. Undetected cases before the first detection do not affect largely because the effect of undetected cases can be compensated by increasing variant frequencies at the dates of their detections. To confirm this, we estimated parameters of the Three-GT-no-I model under assumptions that introduction dates of variants are 10 days and 20 days earlier than the dates of their detections. The estimated values are almost the same as Table 3 except the initial frequencies (Supplementary Table 2 and Supplementary Table 3).

The estimations using the Three-GT-no-I model completely depend on the variant counts based on the nucleotide sequences submitted to the GISAID database from Denmark. Most sequences from Denmark were submitted by member laboratories of Danish Covid-19 Genome Consortium, which analyzed positive specimen from hospitals [23]. As of February 24, the sequencing rate from the 47th week of 2021 to the 7th week of 2022 is 7.2% (157,049 / 2,156,159). Despite of such a high sequencing rate, there may be some extent of sequencing bias toward new variants. Surveillance data that ensure

random sampling of variants are necessary to calculate relative generation times and reproduction numbers accurately.

Acknowledgement

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. World Health Organization, *Tracking SARS-CoV-2 variants*, 2021. Available from: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants

2. World Health Organization, *Weekly epidemiological update on COVID-19 - 1 February 2022*, 2022. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---1-february-2022

3. R. Khandia, S. Singhal, T. Alqahtani, M. A. Kamal, N. A. El-Shall, F. Nainu, et al., Emergence of SARS-CoV-2 Omicron (B.1.1.529) variant, salient features, high global health concerns and strategies to counter it amid ongoing COVID-19 pandemic, *Environ. Res.*, **209** (2022), 112816. https://doi.org/10.1016/j.envres.2022.112816

4. P. A. Desingu, K. Nagarajan, Omicron BA.2 lineage spreads in clusters and is concentrated in Denmark, *J. Med. Virol.*, **94** (2022), 2360–2364. https://doi.org/10.1002/jmv.27659

5. P. J. Halfmann, S. Iida, K. Iwatsuki-Horimoto, T. Maemura, M. Kiso, S. M. Scheaffer, et al., SARS-CoV-2 Omicron virus causes attenuated disease in mice and hamsters, *Nature*, **603** (2022), 687–692. https://doi.org/10.1038/s41586-022-04441-6

6. H. Shuai, J. F. W. Chan, B. J. Hu, Y. Chai, T. T. T. Yuen, F. F. Yin, et al., Attenuated replication and pathogenicity of SARS-CoV-2 B.1.1.529 Omicron, *Nature*, **603** (2022), 693–699. https://doi.org/10.1038/s41586-022-04442-5

7. C. Fraser, Estimating individual and household reproduction numbers in an emerging epidemic, *PLoS One*, **2** (2007), e758. https://doi.org/10.1371/journal.pone.0000758

8. A. Selby, *Estimating Generation Time Of Omicron*, 2022. Available from: http://sonorouschocolate.com/covid19/index.php?title=Estimating_Generation_Time_Of_Omicron

9. J. A. Backer, D. Eggink, S. P. Andeweg, I. K. Veldhuijzen, N. van Maarseveen, K Vermaas, et al., Shorter serial intervals in SARS-CoV-2 cases with Omicron BA.1 variant compared with Delta variant, the Netherlands, 13 to 26 December 2021, *Euro Surveill.*, **27** (2022), 2200042. https://doi.org/10.2807/1560-7917.ES.2022.27.6.2200042

10. K. Leung, M. Lipsitch, K. Y. Yuen, J. T. Wu, Monitoring the fitness of antiviral-resistant influenza strains during an epidemic: a mathematical modelling study, *Lancet. Infect. Dis.*, **17** (2017), 339–347. https://doi.org/10.1016/S1473-3099(16)30465-0

11. K. Leung, M. H. Shum, G. M. Leung, T. T. Lam, J. T. Wu, Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020, *Euro Surveill.*, **26** (2021), 2002106. https://doi.org/10.2807/1560-7917.ES.2020.26.1.2002106

12. K. Ito, C. Piantham, H. Nishiura, Predicted dominance of variant Delta of SARS-CoV-2 before Tokyo Olympic Games, Japan, July 2021, *Euro Surveill.*, **26** (2021), 2100570. https://doi.org/10.2807/1560-7917.ES.2021.26.27.2100570

13. K. Ito, C. Piantham, H. Nishiura, Relative instantaneous reproduction number of Omicron SARS-CoV-2 variant with respect to the Delta variant in Denmark, *J. Med. Virol.*, **94** (2022), 2265–2268. https://doi.org/10.1002/jmv.27560

14. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality, *Euro Surveill.*, **22** (2017), 30494. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494

15. A. Rambaut, E. C. Holmes, Á. O'Toole, V. Hill, J. T. McCrone, C. Ruis, et al., A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology, *Nat. Microbiol.*, **5** (2020), 1403–1407. https://doi.org/10.1038/s41564-020-0770-5

16. Johns Hopkins University, *Johns Hopkins Coronavirus Resource Center*. Available from: https://coronavirus.jhu.edu/

17. W. S. Hart, E. Miller, N. J. Andrews, P. Waight, P. K. Maini, S. Funk, et al., Generation time of the Alpha and Delta SARS-CoV-2 variants, *medRxiv*, (2021), 2021.10.21.21265216. https://doi.org/10.1101/2021.10.21.21265216

18. Y. Pawitan, *In All Likelihood: Statistical Modelling and Inference Using Likelihood*, 2013. Croydon: Oxford University Press.

19. A. R. Conn, N. I. M. Gould, P. Toint, A Globally Convergent Augmented Lagrangian Algorithm for Optimization with General Constraints and Simple Bounds, *SIAM J. Numer. Anal.*, **28** (1991), 545–572. https://doi.org/10.1137/0728030

20. H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control, **19** (1974), 716–723. https://doi.org/10.1109/TAC.1974.1100705

21. A. Cori, N. M. Ferguson, C. Fraser, S. Cauchemez, A new framework and software to estimate time-varying reproduction numbers during epidemics, *Am. J. Epidemiol.*, **178** (2013), 1505–1512. https://doi.org/10.1093/aje/kwt133

22. C. Piantham, K. Ito, Predicting the time course of replacements of SARS-CoV-2 variants using relative reproduction numbers, *medRxiv*, (2022), 2022.03.30.22273218. https://doi.org/10.1101/2022.03.30.22273218

23. Danish Covid-19 Genome Consortium. *Genomic overview of SARS-CoV-2 in Denmark*, 2021. Available from: https://www.covid19genomics.dk/statistics