A new approach to modeling pre-symptomatic incidence and transmission time of imported COVID-19 cases evolving with SARS-CoV-2 variants

Sam Li-Sheng Chen1 · Grace Hsiao-Hsuan Jen1 · Chen-Yang Hsu2 · Amy Ming-Fang Yen1 · Chao-Chih Lai2,3 · Yen-Po Yeh2,4 · Tony Hsiu-Hsi Chen2

Accepted: 24 August 2022 / Published online: 11 September 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

There is paucity of the statistical model that is specified for data on imported COVID-19 cases with the unique global information on infectious properties of SARS-CoV-2 variant different from local outbreak data used for estimating transmission and infectiousness parameters via the established epidemic models. To this end, a new approach with a four-state stochastic model was proposed to formulate these well-established infectious parameters with three new parameters, including the pre-symptomatic incidence rate, the median of pre-symptomatic transmission time (MPTT) to symptomatic state, and the incidence (proportion) of asymptomatic cases using imported COVID-19 data. We fitted the proposed stochastic model to empirical data on imported COVID-19 cases from D614G to Omicron with the corresponding calendar periods according to the classification GISAID information on the evolution of SARS-CoV-2 variant between March 2020 and Jan 2022 in Taiwan. The pre-symptomatic incidence rate was the highest for Omicron followed by Alpha, Delta, and D614G. The MPTT (in days) increased from 3.45 (first period) ~ 4.02 (second period) of D614G until 3.94–4.65 of VOC Alpha but dropped to 3.93–3.49 of Delta and 2 days (only first period) of Omicron. The proportion of asymptomatic cases increased from 29% of D-614G period to 59.2% of Omicron. Modeling data on imported cases across strains of SARS-CoV-2 not only bridges the link between the underlying natural infectious properties elucidated in the previous epidemic models and different disease phenotypes of COVID-19 but also provides precision quarantine and isolation policy for border control in the face of various emerging SRAS-CoV-2 variants globally.

Keywords COVID-19 · Pre-symptomatic · Stochastic process

1 Introduction

It is well acknowledged that transmission and infectiousness parameters such as generation time, series interval, latency, and incubation period varied with SARS-CoV-2 variants according to the previous studies that have already quantified the dynamics of infectious process by estimating these relevant parameters (Ali et al. 2022a; Park et al. 2021; Pung et al. 2021; Campbell et al. 2021). To further assess the spread of SARS-CoV-2 and the elimination of epidemic after containment measures, a series of the previously established epidemic models have been used. The Susceptible-Exposed-Infectious-Removed (SEIR) model is often used to model the basic reproductive number, denoted by R0, for estimating the strength of the spread of community-acquired outbreak of each SARS-CoV2 variant
and effective reproductive number (Rt) for assessing whether the epidemic is controlled according to whether Rt is smaller than 1. Some studies applied the extended SEIR model to estimating the effect of mitigation strategies on the number of infectives and the sequelles of COVID-19 including hospitalization and deaths (Khyar and Allali 2020; Arruda et al. 2021; Gonzalez-Parra et al. 2021). To distinguish the infected persons before onset of symptoms, namely pre-symptomatic or non-persistent asymptomatic cases, from persistent asymptomatic cases, the modified SIR models was also developed to estimate the dynamics of the asymptotically infected people (Sarkar, 2020; Gevertz, 2020). Similar but in-depth SEIR-based models were also adjusted to fit in with the context of asymptomatic infection (Batista et al. 2020; Ahmed et al. 2021; Chatterjee et al. 2021; Ali et al. 2022b; Ma and Lin. 2022; Massard et al. 2022; Yi et al. 2022). While these epidemic models have provided valuable information on transmissibility and infectiousness based on local community-acquired outbreak data before and after the provision of containment measures for each strain of SARS-CoV-2, it is of great interest to model these relevant parameters with the global evolution of a series of SARS-CoV-2 variants from the wild type, D614G, and various kinds of VOCs including Alpha, Beta, Gamma, Delta, and Omicron following periodically tracking the SARS-CoV-2 variants reported by the WHO (WHO, 2022). The SARS-CoV-2 variants including Alpha, Beta, Gamma, and Delta till the recent emergence of Omicron variant have been categorized as Variant of Concern (VOC) based on the risk posed to global public health.

In contrast to local outbreak data based on contact tracing, test, quarantine and isolation, data derived from imported COVID-19 cases have unique but different information on infectious process and containment measures. First, they embrace various kinds of SARS-CoV-2 infection evolving in chronological order from each country or region worldwide. Second, information from imported cases is intractably amenable to estimating transmission and infectiousness parameters if the previous epidemic models are used. For example, it is difficult to know the definite time of exposure to infectives from imported cases but it is available from local outbreak data through contact tracing. Third, imported cases arising from in-bound passenger provide unique information on the states of disease phenotypes including pre-symptomatic and symptomatic phase, namely before and after onset of symptoms, and persistent asymptomatic cases during infectious process when regular RT-PCR test, quarantine, and isolation are requested for in-bound passengers during COVID-19 pandemic. It is therefore interesting to formulate transmission and infectiousness parameters estimated by the previous epidemic models with new parameters of different disease phenotypes expressed by the pre-symptomatic incidence rate, the transition time from pre-symptomatic to symptomatic phase, and the incidence (proportion) of asymptomatic cases estimated form imported COVID-19 data. Implications for these new parameters are several-fold. First, they play an important role in precision testing with reverse transcription polymerase chain reaction (RT-PCR) and also the optimal scheduled interval for quarantining and isolating suspected imported cases. For instance, it would be very interesting to model the transmission time from pre-symptomatic to symptomatic phase as the longer the pre-symptomatic transmission time the longer the interval for quarantine and isolation of suspected imported cases is required. More importantly, as three disease statuses are affected by types of SARS-CoV-2 variants, the policy of non-pharmaceutical intervention (NPI), and the administration of vaccine three of which varied from country to country, the magnitudes of pre-symptomatic incidence rate and the proportion of persistent asymptomatic cases among in-bound passengers would reflect how the underlying risk of infection varied with various strains of SARS-CoV-2 and has been affected by these public health interventions. However, there is paucity of such a kind of statistical model that is tailored for modelling such an unique imported COVID-19 data as mentioned above.

We are therefore motivated by applying the proposed stochastic process to the empirical data on Taiwanese imported cases from various countries and regions worldwide in different periods in parallel with the evolution of SARS-CoV-2 variants, the change in NPI, and the administration of vaccination. Using imported cases gives a clue to the global profiles of SARS-CoV-2 infection and containment measures during COVID-19 pandemic. In addition, our Taiwanese dataset provides a natural opportunity to model natural infectious process and disease progression from pre-symptomatic phase to symptomatic phase, on which we are based to provide evidence-based precision test and scheduled quarantine and isolation.

2 Methods

2.1 The stochastic model for infectious process of the imported COVID-19 case

Figure 1a delineates the continuous-time four-state stochastic model for infectious process of SARS-CoV-2 specified for COVID-19 imported cases. Under the context of stochastic process, the state space is denoted as $\Omega = \{\text{uninfected (state 1), pre-symptomatic phase (PSP) (state 2), symptomatic phase (SP) (state 3), and asymptomatic phase (ASP) (state 4)}\}$. The transitions with
instantaneous potential are allowed from state 1 to state 2 (incidence of PSP), state 1 to state 4 (incidence of ASP), and state 2 to state 3 (transition from PSP to SP).

Let \( X(t) \) denote a random variable for the state at time \( t \) realized by four states Given time \( t \), PSP and ASP belong to symptom-free state but the distinction between PSP and ASP is that the former would progress to SP (non-persistent asymptomatic) whereas the latter would remain free of symptom after infection (persistent asymptomatic). Figure 1b shows the entire infectious process from exposure to infectives, through ASP or PSP until SP for the imported COVID-19 cases accompanied with time to exposure, the arrival time, and the duration of quarantine and isolation of those suspected infectives.

There are several assumptions made for the proposed four-state stochastic process that are supposed to be biologically plausible.

1. The proposed model is progressive from uninfected to SP.
2. ASP would not surface to SP and can only be detected through RT-PCR test.

Note that there is a fraction of patients infected with SARS-CoV-2 will not present any clinical symptoms. We therefore applied the idea of competing risk model to PSP and ASP for those infected. Infected subjects, when going through the pathway of PSP, would have a finite infectiousness time from PSP to SP.

Following the random process of \( X(t) \), Fig. 1b also sketches an illustration of natural course of infection from
departure time until the end of quarantine and isolation. Suppose the \( i \)th enrolled individual had been in close contact with three types of COVID-19 infectives (including symptomatic, pre-symptomatic, and possible asymptomatic cases) (Ferretti et al. 2020) while staying abroad since calendar time \( t_1 \) (departure time). Upon the arrival time \( t_2 \), this individual may have four possible outcomes as defined within \( \Omega \) from the date of exposure to the date of arrival governed by the infectious course of SARS-CoV-2 infection.

We are very interested in estimating three parameters deriving from the above-mentioned four-state stochastic process. These include the incidence rate of pre-symptomatic and asymptomatic COVID-19 and the hazard from pre-symptomatic to symptomatic phase. The parameter of the latter will be converted to the median of pre-symptomatic transmission time (MPTT). The proportion of asymptomatic cases will be also computed by using these parameters.

### 3 Empirical data

To model the infectious process in connection with the four-state disease model of COVID-19, we targeted the imported cases of COVID-19 among inbound passengers around the world flown into Taiwan between March 2020 and Jan 2022 (Table 1). We divided the study period into seven epochs to cover different periods of emerging SARS-CoV-2 variants including the D614G, Variant of Concern (VOC) Alpha, and Delta, and the recent VOC Omicron. The seven epochs were named according to the classification GISAID information on the evolution of SARS-CoV-2 variant, including D614G-1 (March-June 2020), D614G-2 (July–September 2020), Alpha-1 (October-December 2020), Alpha-2 (January-May 2021), Delta-1 (June–August 2021), Delta-2 (September–November 2021), and Omicron-1 (December 2021-January 2022).

It should be noted that we excluded domestic cases because they may contain unknown origin of contact history that may preclude us from estimating the relevant parameters governing the natural history of COVID-19. For each confirmed case, we retrieved data from a repository summarizing the information on imported cases reported by the Central Epidemic Command Center (CECC) in Taiwan (TCDC, 2020). In addition to personal attributes including age and sex, the time stamped on the date of arrival and departure from foreign countries, date of arrival at Taiwan, date of the occurrence of clinical symptoms, and date of RT-PCR test performed were abstracted from the CECC press. It should be noted that as the majority of inbound passengers are Taiwan residents who had been abroad and flown back after business the period between the date of departure and the date of arrival can be exploited for estimating three parameters of interest as below. We can also assume the date of departure would stay in the uninfected state because they would be requested for negative RT-PCR test before departure. Therefore, although the exact date of exposure is unknown it can be assumed the date of exposure must lie between the date of departure and the date of arrival. This forms the basis for building up the likelihood functions as below for estimating pre-symptomatic incidence and the median time of pre-symptomatic transmission. Following the guideline of Taiwan CECC, for subject with suspected symptoms of COVID-19 including fever, cough, short of breath, fatigue, myalgia, diarrhea, and anomaly in smell and test, they would be tested for SARS-CoV-2 infection with RT-PCR twice within 24 h. For inbound passengers without symptoms, it is mandatory for these subjects to be isolated for quarantine for 14 days. The quarantined subjects will be tested for SARS-CoV-2 upon the occurrence of suspected symptoms during their 14-day quarantine (Taiwan Centers for Disease Control 2021). Information on the origin of country/region for the inbound cases were also collected. Supported by the empirical data of timeline through COVID-19 symptom development, we estimated the rates

| Visitors arrivals | Confirmed cases* | Total |
|-------------------|------------------|-------|
|                   | Asymptomatic | Pre-symptomatic | Symptomatic |       |
| D614G-1           | 294,090       | 12 (3.6%)       | 148 (44.0%) | 176 (52.4%) | 338   |
| D614G-2           | 142,015       | 23 (28.4%)      | 15 (18.5%)  | 43 (53.1%)  | 82    |
| Alpha-1           | 146,076       | 202 (65.2%)     | 62 (20.0%)  | 46 (14.8%)  | 310   |
| Alpha-2           | 207,713       | 248 (64.4%)     | 77 (20.0%)  | 60 (15.6%)  | 430   |
| Delta-1           | 86,856        | 191 (70.2%)     | 43 (15.8%)  | 38 (14.0%)  | 276   |
| Delta-2           | 126,473       | 496 (83.1%)     | 50 (8.4%)   | 51 (8.5%)   | 626   |
| Omicron           | 96,894        | 172 (61.2%)     | 54 (19.2%)  | 55 (19.6%)  | 793   |

*593 confirmed cases had censored status of symptoms
of progression from pre-symptomatic phase to symptomatic phase. Data on the number of passengers arrived at Taiwan with the information on the departure countries were retrieved from the open data source of Ministry of Transportation and Communications, Taiwan and Ministry of the Interior National Immigration Agency, Taiwan.

4 Transition probabilities of state transition

For the proposed four-state model, we derived the transition probabilities for various transition modes to develop the total likelihood function for parameter estimation. The transition probabilities from uninfected to pre-symptomatic, symptomatic, and asymptomatic, and of staying in uninfected in the time interval of \((t_1, t_2)\) are derived with the following stochastic integral formula. The transition probability from state 1 (uninfected) to state 2 (PSP) between \(t_1\) and \(t_2\) is expressed as follows,

\[
P_{12}(t_1, t_2) = \int_{t_1}^{t_2} i(s - t_1) \times [1 - F(t_2 - s)] \times [1 - A(s - t_1)] ds
\]  

(1)

where \(s\) refers to the time of being exposed to infectives, which is often unobservable but lies between the time interval of \((t_1, t_2)\), \(i(t)\) represents the probability density function for the transition from state 1 to state 2 at time \(t\), \(F(t)\) is the cumulative distribution for the progression from PSP to SP (transition from state 2 to state 3), and \(A(t)\) is the cumulative distribution of the occurrence of asymptomatic COVID-19.

The first and second element inside the integral of Eq. (1) indicate the process for an infected individual who entered into the PSP at calendar time \(s\) since time \(t_1\) (the date of departure) as shown in Fig. 1 but has not developed into SY yet until time \(t_2\). The third element is not allowed to enter into the ASP once become pre-symptomatic cases according to our model specification.

For the simplest scenario of constant hazards, the Eq. (1) can be expressed as follows,

\[
P_{12}(t_1, t_2) = \int_{t_1}^{t_2} \lambda_1 e^{-\lambda_1 (s-t_1)} \times e^{-\lambda_2 (t_2-s)} \times e^{-\lambda_3 (s-t_1)} ds
\]

where \(\lambda_1, \lambda_3\) represent pre-symptomatic incidence rate, the progression rate from PSP to SP, and asymptomatic incidence rate in the language of epidemiology.

Equation (2) is the transition probability for an individual who had been exposed at time \(t_1\) and entered into PSP at time \(s\) (first element) and turned into symptomatic case at time \(t_2\) (second element). The third element asserts that it would not be possible to go the pathway of being ASP from \(t_1\) to \(t_2\)

\[
P_{13}(t_1, t_2) = \int_{t_1}^{t_2} i(s - t_1) \times [1 - A(s - t_1)] \times F(t_2 - s) ds
\]

(2)

Equation (3) is the transition probability for an individual who become asymptomatic at time \(s\) (first element) and would not follow the pathway of progression from PSP to SP (second element).

\[
P_{14}(t_1, t_2) = \int_{t_1}^{t_2} a(s - t_1) \times [1 - I(s - t_1)] ds
\]

(3)

The complementary probability for an individual staying uninfected state would be derived as follows.

\[
P_{11}(t_1, t_2) = 1 - P_{12}(t_1, t_2) - P_{13}(t_1, t_2) - P_{14}(t_1, t_2)
\]

(4)

The most important information here is related to the probability of progression from PSP to SP in the time interval between arrival \(t_2\) and the end of quarantine \(t_3\),

\[
P_{23}(t_2, t_3) = \int_{t_2}^{t_3} f(r) dr,
\]

which is equivalent to \(F(t_3 - t_2)\).

Equation (5) is to provide the transition probability for a pre-symptomatic individual who develops symptoms at instantaneous time \(r\) during the time interval between \(t_1\) and \(t_2\).

5 Statistical distribution of \(i(t)\) and \(f(t)\)

For the progression rate from PSP to SP in relation to \(f(t)\), various forms can be adapted. According to the distribution of viral load from PSP to SP from literature (Sethuraman et al. 2020), we reckon that the log-logistic form seems more appropriate than others like the Weibull distribution form because it is more likely to have a non-monotonic hazard function to describe the progression from pre-symptomatic to symptomatic phase with time. With the log-logistic form, it is possible to have a hazard function which increases with time when disease progresses and turns to decrease beyond a time point when a patient starts to recover (Collett 2003). The hazard function of log-logistic form is expressed as

\[
h(t) = \frac{e^\theta \kappa t^{\kappa-1}}{1 + e^\theta t^\kappa} \quad 0 \leq t, \kappa > 0
\]

(6)

The cumulative risk of developing symptomatic phase from \(t_1\) to \(t_2\) can be written as follows,

\[
F(t_2, t_3) = 1 - \left(1 + e^{\theta} \cdot (t_2 - t_1)^{\kappa} \right)^{-1}
\]

(7)

Despite this postulate, we attempt a serious of survival functions, including exponential and Weibull distribution, to compare the resulting distribution of emerging
symptomatic cases by time with that of log-logistic form. Recall that the hazard from pre-symptomatic to symptomatic phase will be converted to the median of pre-symptomatic transmission time (MPTT).

6 Parameter estimation with Bayesian Markov Chain Monte Carlo method underpinning

We used the Bayesian Markov Chain Monte Carlo (MCMC) method to estimate the incidence of pre-symptomatic and asymptomatic COVID-19 and the transition from PSP to SP with time following various distributions, including exponential, Weibull, and log-logistic distribution. For each model, the initial 5000 burn-in samples were discarded. Every 20th sample of the following 100,000 iterations was retained and comprised the posterior distribution of parameters of interests. The posterior mean and 95% credible interval of equal tails were reported for all estimates.

7 The simulation study for assessing statistical power

We conducted a simulation study for evaluating the statistical power of the current sample sizes and also various scenarios on seven epochs of imported COVID-19 data. A series of random samples with sizes between 20 and 10,000 from the imported cases in each period were drawn. Each set was used to estimate parameters pertinent to our proposed four-state stochastic model with Bayesian MCMC algorithm. The 95% credible interval was obtained from the posterior distribution of each parameter. Such procedure was repeated 100 times for assessing the proportion that the 95% credible interval covered the corresponding point estimate of each parameter given varying sample size. The algorithm for this simulation study is elaborated below.

Algorithm (Sampling for statistical power)

For $k = 20, 40, 60, 80, 100-1000$ by 100, 5000, 10000

    For $i = 1, 2, \ldots, 100$ Do

    Sampling $x_{it}$ with replacement from data of confirmed cases

    MCMC sampling for the posterior distribution of parameters $\theta_{ik}$ (intercepts and regression coefficients of $\lambda_1, \lambda_3, \text{theta}_2, r_2$, take log-logistic distribution for the MPTT as an example)

    Let $l_{0_{ik}} = 1$ if $\hat{\theta}_{ik}$ in the current study $\in$ 95% CI of $\theta_{ik}$

    Else $l_{0_{ik}} = 0$

End

Report the statistical power for sample size $k$ with $\frac{\sum_{i=1}^{100} l_{0_{ik}}}{100}$

End
8 Results

Table 2 shows the estimated results of various four-state stochastic models with three corresponding distributions of the infectiousness time from pre-symptomatic phase (PSP) to symptomatic phase (SP). The estimated incidence of pre-symptomatic and asymptomatic cases was 131 (95% CI: 124, 139) and 122 (95% CI: 115, 130) per 100,000, respectively. For the transition from PSP to SP, the scale parameter of the exponential model was estimated as 0.2217 (95% CI: 0.2034, 0.2412), yielding 3.13 days (95% CI: 2.87, 3.41) of median of pre-symptomatic transmission time (MPTT) from PSP to SP (Fig. 2). For the log-logistic model, the logit scale and shape parameters were estimated as -2.8675 (-3.1713, -2.5656) and 2.3420 (95% CI: 2.1590, 2.5295), leading to 3.40 days (95% CI: 3.15, 3.66) of MPTT. The corresponding parameters for the Weibull form were 0.1405 (95% CI: 0.1152, 0.1692) and 1.2173 (95% CI: 1.1383, 1.2958), which gave 3.73 days (95% CI: 3.40, 4.06) of MPTT.

Figure 2 shows the hazard function by time since exposure for these three models. The log-logistic form revealed an increasing hazard in the first four days since exposure in contrast to the constant hazard for the exponential form or the monotonic increasing hazard for the Weibull form.

The results of the DIC statistics show that log-logistic form for fitting the pre-symptomatic infectious time had the lowest DIC (33,597.59), followed by the Weibull distribution (33,911.49) and the exponential distribution (33,940.65) (Table 3). The model allowing for the period effect led to a smaller DIC given each of distribution, of which log-logistic still had the smallest DIC (31,379.62). When we further considered both effects of period and area, all the DICs were further reduced. Again, the log-logistic form still had the smallest DIC (29,255.44).

Figure 3 shows the heat map of the posterior mean of pre-symptomatic and asymptomatic incidence (per 1000 person-year), the proportion of asymptomatic cases, and the MPTT with adjustment for the period effect, including the first D614G period, the second D614G period, the first Alpha period, the second Alpha period, the first Delta period, the second Delta period, and the Omicron period. It is very interesting to note that three surges of the incidence rate of being PSP (Fig. 3a) fit in with the evolution of the SARS-CoV-2 variants, namely, the old type dominated by the emerging variant. It can be clearly seen that the transition from PSP to SP accelerated when the emerging variant pre-dominated whereas the corresponding transition slowed down with a longer MPTT before the next emerging type (Fig. 3b). The MPTT of the first D614G (3.45 days) was shorter than that of the second (4.02 days). This was also noted for VOC Alpha and Delta, while Omicron had the shortest MPTT (2.03 days) compared with other SARS-CoV-2 variants. The Omicron was still rampant in the first period and would be expected to have the same pattern with a longer dwelling time afterwards. It is very interesting to see the proportion of asymptomatic increased with periods (Fig. 3c) before Omicron emerged. It could be attributed to the change of border control policy upon the request of negative test three days before embarking at airport. The proportion of asymptomatic cases after emerging Omicron VOC was 60% while both

| Transition rate          | Exponential model | Log logistic model | Weibull model |
|--------------------------|-------------------|-------------------|---------------|
|                          | Estimate 95% CI   | Estimate 95% CI   | Estimate 95% CI |
| State 1 → State 2, λ₃   | 0.00131 0.00124   | 0.00133 0.00125   | 0.00131 0.00124 |
| State 2 → State 3, λ₂(ₜ)| 0.2217 0.2034     | -2.8675 -3.1713   | 0.1405 0.1152 |
| Scale parameter          | NA                | 2.3420 2.1590     | 2.1713 1.1383 |
| Shape parameter          | 0.00122 0.00115   | 0.00118 0.00110   | 0.00122 0.00115 |
| State 1 → State 4, λ₃   | 0.00139           | 0.00141           | 0.00139 |

Fig. 2 Time-varying hazard function with three distributions of pre-symptomatic transmission time.
Table 3 The DIC statistics for the four-state stochastic models with and without the covariates of period and area given three distributions of pre-symptomatic transmission time

| Model                                                                 | Exponential     | Log logistic     | Weibull      |
|----------------------------------------------------------------------|-----------------|-----------------|--------------|
| Four-state stochastic model                                          | 33,940.05       | 33,597.59       | 33,911.49    |
| Four-state stochastic model regressing on period                     | 31,907.22       | 31,379.62       | 31,847.16    |
| Four-state stochastic model regressing on period and area with a shared area effect in the second D614G period | 29,747.89       | 29,277.98       | 30,048.61    |
| Four-state stochastic model regressing on period and area             | 29,718.81       | 29,255.44       | 29,706.91    |

Fig. 3 The posterior mean of incidence of pre-symptomatic cases, median of pre-symptomatic transmission time, and the proportion and incidence of asymptomatic case by period
pre-symptomatic incidence and asymptomatic incidence were the highest during the study period.

Figure 4 shows the proportion of the 95% CI covering the point estimate in the current study in the simulation with varying size of random sample. It shows that sample size of 200 had sufficient statistical power (> 80%) for the incidence rate of both pre-symptomatic and asymptomatic cases. However, it requires a sample size of 500 in each period to reach sufficient statistical power (> 80%) for the inference of MPTT and the proportion of asymptomatic cases. However, a huge sample size still cannot help the inference of the proportion of asymptomatic cases in the first period of D614G-1 because of sparse events. Given the fact that the event numbers were larger than 200 in each epoch in this study, except only D614G-1 with sparse asymptomatic cases, the statistical power would be sufficient enough for other estimated parameters.

9 Discussion

While a series of established statistical models have been already used to estimate transmission and infectiousness parameters of SARS-CoV-2 infection based on local outbreak data, modelling imported COVID-19 data as proposed here by a four-state stochastic process may bridge the link between these infectious parameters estimated before with the disease phenotypes of infectious process formulated by pre-symptomatic incidence rate, the median of pre-symptomatic transmission time (MPTT), and the incidence (proportion) of asymptomatic cases. So doing is particularly valuable when the proposed model was applied to the empirical data on imported COVID-19 cases covering different SARS-CoV-2 variants in chronological order during COVID-19 pandemic.

Implications for modelling the dynamics of disease axis using imported COVID-19 data are several-fold. First, the magnitude of pre-symptomatic incidence may be a
reflection of transmission coefficient affected by the latency of virus replication and partial information on the generation time between infector and infectee when using the previous epidemic model based on local outbreak data with main emphasis on the axis of infection that starts from exposure to infection and further becomes infectives for infecting other susceptible persons. Higher pre-symptomatic incidence rate estimated by our new approach was in accordance with higher transmission coefficients already demonstrated by the epidemic SEIR model. Higher pre-symptomatic incidence rate in Omicron from our study is consistent with higher infection rate of Omicron demonstrated in the previous studies with higher growth rate and effective reproductive number (Elliott et al. 2022; Ito et al. 2022) because the emergence of Omicron VOC has replaced Delta VOC as the dominant circulating strain worldwide since the end of 2021 (Elliott et al. 2022; Pulliam et al. 2022). Higher transmission coefficient of Omicron VOC may be attributed to both the high viral shedding (Zheng et al. 2022) and immune escape (Pulliam et al. 2022; Chaguza et al. 2022) of Omicron VOC. Again, our estimated results in the remarkable high pre-symptomatic incidence of Omicron VOC is in line with these findings.

Second, the duration of MPTT may provide complementary information with the generation time and the incubation time before the presence symptoms as demonstrated in the previous studies (Ali et al. 2022a; Park et al. 2021; Pung and Mak 2021; Hart et al. 2022; Abbott et al. 2022) although our estimated MPTT may not be directly compared with both as the definition is still different. While the generation time or serial interval estimates varied with SARS-CoV-2 variant the estimates of MPTT also varied with SARS-CoV-2 variant. Hart et al. reported the shorter generation time for the Delta variant (4.7 days) than the Alpha variant (5.5 days) (Hart et al. 2022). Kremer et al. estimated the shorter serial interval (2.75 days) for the Omicron variant than the Delta variant (3.00 day) (Kremer et al. 2022). Our MPTT estimates by the different variants are in line with both estimates of generation time and series intervals reported from previous studies.

Third, the separation of pre-symptomatic cases from asymptomatic case is to avoid the confusion of conventional use in defining asymptomatic cases in static property that actually consist of pre-symptomatic cases and asymptomatic cases at certain time point of infection axis. To this end, the previous compartment SEIR models (Batista et al. 2020; Ahmed et al. 2021; Chatterjee et al. 2021; Ali et al. 2022b; Ma and Lin 2022; Massard et al. 2022; Yi et al. 2022) have been accommodated to estimate the asymptomatic growth and the fraction of asymptomatic infectives was parameterized in the model. Our proposed four-state stochastic process for estimating the proportion (incidence) of asymptomatic cases is tailored for such a purpose in a similar fashion. It is very interesting to note that the reciprocal relationship between the incidence rate of pre-symptomatic cases and asymptomatic cases by periods gives the severity of each variant and the administration of vaccine for reducing severity. The higher prevalence of asymptomatic infection (23%) in Omicron strain compared with other variants was found in South Africa (Garrett et al. 2022). We also found the high proportion of asymptomatic infection in both Delta and Omicron periods. Asymptomatic infection caused by the variants or people with immunity following infection or vaccination should be further studied.

Fourth, the log-logistic form suggests an increasing hazard for symptomatic phase in the initial days after being infected but decreasing after the peak time. From the viewpoint of biological plausibility, this is consistent with the evidence that viral shedding changed over time since exposure (Sethuraman et al. 2020). The results of short Delta and Omicron VOC compared with Alpha VOC are consistent with the lower generation time in Delta VOC (Hart et al. 2022) and Omicron (Abbott et al. 2022).

Quantifying incidence of pre-symptomatic cases and MPTT also provides the possibility of translating the changes in these parameters into the effectiveness of containment measures given each variant. Evaluation of these parameters in chronological order is indicative of the effectiveness of containment measures including NPI, testing, and vaccination varying with period in parallel with the evolution of SARS-CoV-2 variants. Such an application of the proposed model for developing the surveillance in border control of imported cases has been demonstrated by using Taiwanese data on imported cases which covered a range of SARS-CoV-2 variants from Alpha to Omicron between March 2020 and January 2022. The estimated results show that the incidence of the pre-symptomatic cases was aligned with the epidemic curve globally, starting from the initial 109 per 100,000 passengers in the first D614G epoch, dropping to 40 per 100,000 in the second D614G epoch due to the restriction of NPI, resurging to 163 when Alpha VOC preponderated, and decreasing to 117 in the second Alpha epoch. Both findings revealed 63% and 28% effectiveness of NPIs for D614G and Alpha VOC, respectively, before vaccination. The incidence rates of pre-symptomatic phase when vaccine is available were similar between two epochs of the Delta period. However, it is contrary to the expectation of higher incidence of pre-symptomatic cases for Delta compared with Alpha because the former should have higher transmission than the latter in the light of the estimated basic reproductive number (R₀) reported in the previous studies (Campbell et al. 2021; Liu and Rocklöv 2021). The decline in the absolute incidence rate of Delta compared with that of Alpha indicated the possible effectiveness of vaccine in
reducing the frequencies of contact when vaccine reduces the susceptible population and increases immunized population and also reduces the chance of transmission probability. The incidence rate resurred dramatically in the recent Omicron period. This is partially due to the lifting of all NPIs after full vaccination and partially due to the waning of vaccine after full vaccination. More importantly, the incidence of asymptomatic cases increased consistently when time went by from first D614G epoch to Omicron epoch. This is mainly caused by the advent of vaccine that renders asymptomatic cases become more likely and is also possible due to different transmissibility and the severe property of SARS-CoV-2 variants. For example, Omicron had high transmissibility but most of infective are milder than Alpha and Delta.

The estimates of MPTT by SARS-CoV-2 variants revealed the speed of infection pertaining to pre-symptomatic transmission for different kinds of variants. It is very interesting to note that Omicron compared with other variants had higher growth rate of infection that is also commensurate with higher strength of infection in relation to pre-symptomatic transmission. In this sense, the proposed stochastic process may not only capture the feature of the strength-based reproductive number (R_0) model but also accommodate the speed-based growth rate model, which also provide important information on the growth rate of natural infection process as emphasized by Anderson et al. (2020). More importantly, it should be also noted that the estimated MPTT from pre-symptomatic to symptomatic phase by variants has significant implications for providing the optimal length (day) of quarantine and isolation. Based on this finding the length of quarantine can be reduced from Alpha, Delta and Omicron.

The major limitation of this study is pertaining to the generalizability of the proposed methodology. As Taiwan COVID-19 epidemic is well controlled the empirical data based on community-acquired outbreak data across SARS-CoV-2 variants are hardly available. This accounts for why we exploited COVID-19 imported cases for estimating relevant parameters when using the proposed four-state stochastic model. However, whether the findings obtained from imported COVID-19 cases in the current study are similar to those based on community-acquired outbreak data is still uncertain and is worthy of being investigated in the future study. The other aspect of generalizability is relevant to whether Taiwan COVID-19 imported cases can be representative of each variant worldwide. Such an external validation can be made if the proposed four-state stochastic model can be applied to the imported COVID-19 cases of other countries. Finally, our simulation study on statistical power evaluation suggests that modelling imported COVID-19 data is still subject to the number of imported COVID-19 cases. In our study, statistical inference for asymptomatic cases was underpowered in the first epoch in March-June 2020 given sparse events.

In conclusion, modeling data on imported cases across strains of SARS-CoV-2 not only bridges the link between the underlying natural infectious properties elucidated in the previous epidemic models and different disease phenotypes of COVID-19 but also provides precision quarantine and isolation policy for border control in the face of various emerging SRAS-CoV-2 variants globally.

Author’s contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by GHJ, CYH and AMY. The first draft of the manuscript was written by SLSC, CCL, YPY, and THC interpreted results. All authors provided input on the revision of the manuscript. All authors read and approved the final manuscript.

Funding This study was funded by National Science and Technology Council, Taiwan (MOST 108-2118-M-038 -002-MY3; MOST 109-2327-B-002-009; MOST 109–2327-B-002-009). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Abbott S, Sherratt K, Gerstung M, Funk S (2022) Estimation of the test to test distribution as a proxy for generation interval distribution for the omicron variant in England. medRxiv. Doi: https://doi.org/10.1101/2022.01.08.22268920

Ahmed I, Modu GU, Yusuf A, Kumam P, Yusuf I (2021) A mathematical model of coronavirus disease (COVID-19) containing asymptomatic and symptomatic classes. Res Phys 21:103776. https://doi.org/10.1016/j.rinp.2020.103776

Ali ST, Yeung A, Shan S, Wang L, Gao H, Du Z, Xu XK, Wu P, Lau E, Cowling BJ (2022a) Serial Intervals and Case Isolation delays for coronavirus disease 2019: a systematic review and meta-analysis. Clin Infect Dis Official Public Infect Dis Soc Am 74(4):685–694. https://doi.org/10.1093/cid/ciaa491

Ali Z, Rabiei F, Rashidi MM, Khodadadi T (2022b) A fractional-order mathematical model for COVID-19 outbreak with the effect of symptomatic and asymptomatic transmissions. Eur Phys J plus 137(3):395. https://doi.org/10.1140/epjp/s13360-022-02603-z

Anderson R, Donnelly C, Hollingsworth TD, et al (2020) Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation. In: Report for The Royal Society Science in Emergencies Tasking—COVID-19 (SET-C) for communication to the Scientific Advisory Group for Emergencies (SAGE).

Arruda EF, Das SS, Dias CM, Pastore DH (2021) Modelling and optimal control of multi strain epidemics, with application to...
COVID-19. PLoS ONE 16(9):e0257512. https://doi.org/10.1371/journal.pone.0257512
Batista B, Dickenson D, Gurski K, Kebe M, Rankin N (2020) Minimizing disease spread on a quarantined cruise ship: a model of COVID-19 with asymptomatic infections. Math Biosci 329:108442. https://doi.org/10.1016/j.mbs.2020.108442
Bicher M, Rippinger C, Schneckenreither G, Weirecht N, Uraeh C, Zechmeister M, Brunneir D, Huf W, Popper N (2022) Model based estimation of the SARS-CoV-2 immunization level in Austria and consequences for herd immunity effects. Sci Rep 12(1):2872. https://doi.org/10.1038/s41598-022-06771-x
Campbell F, Archer B, Laurenson-Schafer H et al (2021) Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Euro Surveill 26(24):2100509. https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509
Chaguza C, Coppi A, Earnest R et al (2022) Rapid emergence of SARS-CoV-2 omicron variant is associated with an infection advantage over delta in vaccinated persons. Med 3:235–334. https://doi.org/10.1016/j.med.2022.03.010
Chatterjee S, Sarkar A, Karmakar M, Chatterjee S, Paul R. (2021). A mathematical modeling approach. medRxiv. Doi: https://doi.org/10.1101/2021.02.24.21252406.
Collett D (2003) Modeling Survival Data in Medical Research, 2nd edn. Chapman and Hall, London
Elliott P, Eales O, Steyn N et al. (2022). Twin peaks: The Omicron SARS-COV-2 BA.1 and BA.2 epidemics in England. Science, 376(6600):eabq4411. Doi: https://doi.org/10.1126/science. abq4411
Ito K, Plantham C, Nishiura H.(2022) Estimating relative generation times and reproduction numbers of Omicron BA 1 and BA 2 with respect to Delta variant in Denmark. Math Biosci Eng 19(9):9005–17. Doi: https://doi.org/10.3934/mbe.2022418
Ferretti L, Wymant C, Kendall M, et al (2020) Quantifying SARS - CoV-2 transmission suggests epidemic control with digital contact tracing. Science 368:eabb6936. Doi: https://doi.org/10.1126/science.abd6936.
Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, Hens N (2020) Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Euro Surveill 25(17):2000257. https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257
Garrett N, Tapley A, Andriesen J, et al (2022) High Asymptomatic Carriage with the Omicron Variant in South Africa. Clin Infect Dis. ciac237. Doi: https://doi.org/10.1093/cid/ciac237
Geyer JL, Greene JM, Sanchez-Tapia CH, Sontag ED (2021) A novel COVID-19 epidemiological model with explicit susceptible and asymptomatic isolation compartments reveals unexpected consequences of timing social distancing. J Theor Biol 510:110539. https://doi.org/10.1016/j.jtbi.2020.110539
Gonzalez-Parra G, Martinez-Rodriguez D, Villanueva-Mico R. (2021) Impact of a new SARS-CoV-2 variant on the population: A mathematical modeling approach. medRxiv. Doi: https://doi.org/10.1101/2021.02.24.21252406.
Hart WS, Miller E, Andrews NJ, Wright P, Maini PK, Funk S, Thompson RN (2022) Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. Lancet Infect Dis S1473–3099(22):00001–00009. https://doi.org/10.1016/S1473-3099(22)00001-9
Khyar O, Allali K. (2020). Global dynamics of a multi-strain SEIR epidemic model with general incidence rates: application to COVID-19 pandemic. Nonlinear dynamics, pp 1–21. Advance online publication. Doi: https://doi.org/10.1007/s11071-020-05929-4
Kremer C, Braeye T, Proesmans K, Andre E, Torneri A, Hens N. (2022) Serial intervals for SARS-CoV-2 Omicron and Delta variants, Belgium, November 19-December 31, 2021. Emerg Infect Dis 28(8), Doi: https://doi.org/10.3201/eid2808.220220
Liu Y, Rocklov J.(2021) The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. J Travel Med 28(7):taab124. Doi: https://doi.org/10.1093/jtm/taab124.
Ma J, Lin W (2022) Dynamics of a stochastic COVID-19 epidemic model considering asymptomatic and isolated infected individuals. Math Biosci Eng MBE 19(5):5169–5189. https://doi.org/10.3934/mbe.2022242
Massard M, Eftimie R, Perasso A, Saussereau B (2022) A multi-strain epidemic model for COVID-19 with infected and asymptomatic cases: application to French data. J Theor Biol 545:111117. https://doi.org/10.1016/j.jtbi.2022.111117
National Immigration Agency, Ministry of the Interior, Taiwan. https://www.immigration.gov.tw/5475/.
Park SW, Bolker BM, Funk S (2021) Roles of generation-interval distributions in shaping relative epidemic speed, strength, and control of new SARS-CoV-2 variants. medRxiv. Doi: https://doi.org/10.1101/2021.05.03.21256545
Pulliam JR, van Schalkwyk C, Govender N et al. (2022) Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa. Science. 376(6593):eabn4947. Doi: https://doi.org/10.1126/science.abn4947
Pung R, Mak TM; CMMID COVID-19 working group, Kucharski AJ, Lee VJ (2021) Serial intervals in SARS-CoV-2 B.1.617.2 variant cases. Lancet 398(10303):837–838. Doi: https://doi.org/10.1016/ s0140-6736(21)01697-4
Sarkar K, Khajanchi S, Nieto JJ (2020) Modeling and forecasting the COVID-19 pandemic in India. Chaos Solitons Fractals 139:109049. https://doi.org/10.1016/j.chaos.2020.109049
Sethuraman N, Jeremiah SS, Ryo A (2020) Interpreting diagnostic tests for SARS-CoV-2 JAMA. 323:2249–2251. Doi: https://doi.org/10.1001/jama.2020.8259
TCDC, Taiwan Centers for Disease Control. https://www.cdc.gov.tw/En.
Ministry of Transportation and Communications, Taiwan. https://www.motc.gov.tw/en/index.jsp.
World Health Organization, Tracking SARS-CoV-2 variants. Access on 25 June 2022. https://www.who.int/activities/tracking-SARS-CoV-2-variants
Yi GY, Hu P, He W (2022) Characterizing the COVID-19 dynamics with a new epidemic model: susceptible-exposed-asymptomatic-symptomatic-active-removed. Can J Stat 50(2):395–416. https://doi.org/10.1002/cjs.11698
Zheng J, Wang Z, Li J et al (2022) High amounts of SARS-CoV-2 in aerosols exhaled by patients with Omicron variant infection. J Infect 84(6):e126–e128. https://doi.org/10.1016/j.jinf.2022.02.015

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.