1 2	Insights into COVID-19 epidemiology and control from temporal changes in serial interval distributions in Hong Kong
3	
4	Sheikh Taslim Ali <sup>1,2,†</sup> , Dongxuan Chen <sup>1,2,†</sup> , Wey Wen Lim <sup>1,2</sup> , Amy Yeung <sup>1,2</sup> , Dillon C.
5	Adam <sup>1,2</sup> , Yiu Chung Lau <sup>1,2</sup> , Eric H. Y. Lau <sup>1,2</sup> , Jessica Y. Wong <sup>1</sup> , Jingyi Xiao <sup>1</sup> , Faith Ho <sup>1</sup> ,
6	Huizhi Gao <sup>1</sup> , Lin Wang <sup>3</sup> , Xiao-Ke Xu <sup>4</sup> , Zhanwei Du <sup>1,2,</sup> , Peng Wu <sup>1,2</sup> , Gabriel M. Leung <sup>1,2,‡</sup> ,
7	Benjamin J. Cowling <sup>1,2,‡</sup>
8	
9	Affiliations:
10	<sup>1</sup> WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of
11	Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong
12	Special Administrative Region, China
13	<sup>2</sup> Laboratory of Data Discovery for Health Limited, Hong Kong Science and Technology Park,
14	New Territories, Hong Kong.
15	<sup>3</sup> Department of Genetics, University of Cambridge, Cambridge, United Kingdom,
16	<sup>4</sup> College of Information and Communication Engineering, Dalian Minzu University, Dalian
17	116600, China.
18	
19	<sup>†</sup> These authors contributed equally to this work.
20	<sup>‡</sup> Joint senior authors
21	
22	Corresponding author: bcowling@hku.hk
23	
24	
25 26	
27	
28	
29 20	
30 31	
32	NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

3	3
-	-

## 34 Abstract:

35

36 The serial interval distribution is used to approximate the generation time distribution, an essential parameter to predict the effective reproductive number " $R_t$ ", a measure of 37 38 transmissibility. However, serial interval distributions may change as an epidemic progresses 39 rather than remaining constant. Here we show that serial intervals in Hong Kong varied over 40 time, closely associated with the temporal variation in COVID-19 case profiles and public 41 health and social measures that were implemented in response to surges in community 42 transmission. Quantification of the variation over time in serial intervals led to improved estimation of  $R_t$ , and provided additional insights into the impact of public health measures on 43 44 transmission of infections.

- 45
- 46

## 47 **One-Sentence Summary:**

48

49 Real-time estimates of serial interval distributions can improve assessment of COVID-19

50 transmission dynamics and control.

## 51 Main Text:

52 Monitoring the intensity of coronavirus disease 2019 (COVID-19) transmission is an essential 53 component of public health surveillance for situational awareness and real-time impact 54 assessment of interventions (1, 2). Transmissibility has typically been measured by the effective reproductive number " $R_t$ " based on analysis of epidemic curves along with an 55 estimate of the generation time distribution (3-5). The generation time for an infectious disease 56 57 describes the average time between consecutive infections in a transmission chain (4, 6), and 58 is often approximated by the serial interval distribution which describes the average time 59 between illness onsets of consecutive cases in a transmission chain (3, 7, 8). However, we and 60 others have shown that the serial interval distribution may change as an epidemic progresses 61 rather than remaining constant (9-11).

62

In Hong Kong, a subtropical city located on the southern coast of China, the public health 63 responses to COVID-19 following a "dynamic zero covid" strategy aiming to eliminate local 64 65 infections included a number of key components. First, travel restrictions and on-arrival quarantines on all in-bound travellers have been implemented throughout the pandemic to 66 67 reduce the importation risk. Second, a series of targeted and community-wide public health and 68 social measures have been used to minimize local transmission of COVID-19, including 69 hospital isolation of all confirmed cases, quarantine of close contacts in designated facilities, 70 mask mandate, restriction on gatherings, closures of facilities, schools, and some workplaces, 71 and compulsory testing orders for persons at a higher risk of exposure. This strategy was 72 successful in limiting cumulative incidence to 12,600 confirmed cases by 31 December 2021, 73 corresponding to 0.16% of the total population in Hong Kong. All community epidemics were 74 associated with ancestral-like viruses and not variants. Here, our objective is to take advantage of detailed contact tracing data in Hong Kong to characterise temporal changes within and 75

between waves in the serial interval distribution of COVID-19 through a series of community
epidemics, and identify possible factors associated with those temporal changes.

- 78
- 79

## 80 COVID-19 transmission and PHSMs in Hong Kong

81 We obtained detailed information on each laboratory-confirmed case of COVID-19 from the Department of Health of the Hong Kong SAR government (section 1, supplementary 82 83 materials). The database included information on the confirmation date, illness onset date, 84 isolation date, hospital admission date, outcome (critical/serious/stable), infection origin 85 (locally infected vs infected outside Hong Kong) and arrival date (if applicable), home and 86 workplace location, detailed travel/movement history, cluster information from contact tracing 87 (if applicable) of each confirmed case (12). In Hong Kong from 22 January 2020 through 31 July 2022, COVID-19 has caused five local epidemic waves. The "first wave", defined as the 88 89 period from late January to mid-February 2020 with a small number of community cases linked 90 to importations from mainland China (13). Given the very small number of imported infections 91 (n=45) we excluded the first wave from our study. We analysed data from throughout the 92 second wave (1 March 2020 – 10 April 2020), third wave (25 June 2020 – 8 September 2020) 93 and fourth wave (1 Nov 2020 – 23 March 2021) (Fig. 1A). In the fifth wave which started in 94 early January 2022 (14), contact tracing capacity was challenged due to the fast exponential 95 rise in cases, and we were only able to include detailed data on confirmed transmission pairs from the very early stage of this wave in these analyses. Our study received ethical approval 96 97 from the Institutional Review Board of the University of Hong Kong (ref: UW 20-341).

98

We collected information from official reports on the public health and social measures(PHSMs) implemented in Hong Kong to control the spread of COVID-19 (*15*). We further

101 classified these interventions into case-based, community-wide (including travel-based) 102 control measures (*16*). The case-based measures including strict isolation of cases and 103 quarantine of close contacts were maintained throughout each wave while the community-wide 104 measures were generally only implemented when needed to bring each epidemic wave under 105 control. The summary of these PHSMs including their timing and duration is provided in 106 Supplementary table S1.

107

## 108 **Construction of transmission pairs**

109 Following our earlier work (17), we first reconstructed the initial transmission pairs with 110 reference to the detailed individual data on all confirmed cases retrieved from the Centre for 111 Health Protection. We further rechecked the initially constructed pairs by their cluster 112 information and epidemiological linkage with other cases and determined the infector and 113 infectee within a pair according to a pre-defined algorithm (section 1, supplementary 114 materials). The larger clusters with complex epidemiological linkages between the cases were 115 individually assessed to determine the infector and the infectee in each transmission pair. If 116 there were two or more likely infectors in one cluster, we defined the infector as the case with 117 the earliest onset date. If cases shared the same onset date, the one with an earlier report date or case number would be classified as an infector (section 1, supplementary materials) (9, 18). 118 119 We also performed a possible cross-check on unclear pairs with phylogenetic data to examine 120 potential infector-infectee relationships with the *Phybreak* package in R (19). We resolved the 121 multiple infector issue by identifying a case to be the infector of all subsequent cases with an 122 immediate link in a cluster when the onset intervals between the infector and the infectees fell 123 within a pre-defined time period, 8 days for the main analysis and 5-14 days for the sensitivity 124 analyses (section 6, supplementary materials). Among more than 12,600 COVID-19 cases in 125 Hong Kong in the period 1 March 2020 through to 23 March 2021, most were identified in the third and fourth waves (Fig. 1A), in total we were able to construct 2433 transmission pairs,
including 87, 965 and 1381 in the second, third and fourth waves, respectively. Among these,
we considered 47, 357 and 355 transmission pairs as "confirmed pairs" in these three waves,
respectively, and the remainder were "likely pairs" (section 1, supplementary materials). There
were an additional 64 pairs identified in small outbreaks occurred between our pre-defined
waves, and 229 pairs (30 for BA1, 174 for BA2 and 25 for delta) identified in the very early
stage of the fifth wave (fig. S1).

133

## 134 Time-varying serial intervals and PHSMs

We estimated the serial interval distributions over time in a Bayesian framework based on 135 136 Markov Chain Monte Carlo, implemented with the RStan package in R (section 2, 137 supplementary materials). We estimated serial interval distributions for three periods within 138 each wave i.e. pre-peak, during peak, and post-peak, and then estimated the time-varying daily 139 effective serial interval distributions with a sliding window of 10 days (7-14 days for sensitivity 140 analysis) (9). We used the same framework to estimate the distributions of onset-to-isolation 141 intervals (indicating case isolation delay) for the infectors and then performed multivariable 142 regression analysis on mean serial intervals where the mean onset-to-isolation interval and 143 various PHSMs were included as explanatory variables (section 3, supplementary materials). 144 The variation in population mobility and mixing, potentially reflecting the effect of these 145 PHSMs, might also impact the transmission dynamics of COVID-19 (20-22). Therefore, we also performed a multivariable regression analysis on the estimated mean serial intervals with 146 147 the daily digital transactions through Octopus cards, a widely-used contactless electronic 148 payment system in Hong Kong, as the proxy of relative mobility of the population (20), and 149 daily per capita testing volume (number of PCR tests conducted per 10,000 population per day) 150 in Hong Kong (section 3, supplementary materials).

151

Among the 2497 transmission pairs identified in during the first two years of the COVID-19 pandemic in Hong Kong including the second, third and fourth waves with the ancestral strain of SARS-CoV-2, the mean serial interval ( $\mu$ ) was estimated to be 3.6 (95% CrI: 3.5, 3.7) days with standard deviation ( $\sigma$ ) 3.4 (95% CrI: 3.3, 3.5) days.

156

157 We identified the peak of the second wave from 16 to 24 March 2020, from 18 to 27 July 2020 158 for the third wave, and the fourth wave had two peaks, the first from 30 November 2020 to 15 159 December 2020 and the second during 11 - 25 January 2021 (section 1, supplementary 160 materials). There were clear temporal changes in mean serial interval estimates within each 161 epidemic wave studied and across waves, with mean serial intervals shortening from 5.5 days (95% CrI: 4.4, 6.6) for pre-peak to 3.2 days (95% CrI: 1.9, 4.4) for post-peak in the second 162 163 wave, and from 4.6 (95% CrI: 4.1, 5.0) days to 2.7 (95% CrI: 2.2, 3.2) days in the third wave 164 (Fig. 1A and table S2). However, less clear changes were identified in the fourth wave where 165 two peaks in incidence were noted (Fig. 1A). The standard deviations of serial interval 166 distributions within individual epidemic waves were more comparable but inter-wave 167 differences were noticeable ranging from 2.4 days to 4.3 days (table S2). The mean onset-to-168 isolation intervals had a similar decreasing temporal pattern along with mean serial intervals 169 during the second wave shortening from 5.8 days to 3.5 days, while the estimates were largely constant during the third wave between 5.2 days and 5.4 days and fluctuated throughout the 170 171 fourth wave (3.4 days to 5.0 days) (Fig. 1A and table S2). In the early part of the fifth wave, the mean serial interval was estimated to be 3.6 (95% CrI: 3.5, 3.7) days with standard deviation 172 173 3.4 (95% CrI: 3.3, 3.5) days with comparable estimates for the Delta variant ( $\mu$ : 4.0 days,  $\sigma$ : 2.4 days), the Omicron BA.1 ( $\mu$ : 3.3 days,  $\sigma$ : 2.0 days) and the Omicron BA.2 ( $\mu$ : 3.6 days,  $\sigma$ : 174 175 1.8 days) subvariants.

176

177 The time-varying mean effective serial intervals and onset-to-isolation intervals across the 178 second, third and fourth epidemic waves in Hong Kong showed similar temporal patterns in 179 the two epidemiological parameters based on the daily changes in the estimates (Fig. 1, B to 180 **D**). The daily variations (shortening and lengthening) in the estimates of two parameters either 181 occurred towards the same direction (most of the second wave, some phases of the third and 182 the fourth waves) or opposite directions (some of the phases in the third wave and the fourth 183 wave) over the epidemics. Similar results were also indicated in the sensitivity analysis on all 184 transmission pairs (fig. S2).

185

186 The implemented PHSMs could explain up to 70%, 42% and 49% of the variance in the mean 187 effective serial intervals estimated for the second, third and fourth waves, respectively (table 188 S3). The onset-to-isolation interval, reflecting the impact of the targeted measure of case 189 identification and isolation, was found to have statistically significant positive associations 190 with the effective serial intervals for the second and third waves with up to 60% and 13% of 191 the variations explained in the estimated effective serial intervals for the two waves 192 respectively, but there was not a statistically significant association in the fourth wave. The 193 interrupted regression further suggested similar associations between onset-to-isolation 194 intervals and effective serial intervals during different phases of the epidemics (table S4). A 195 maximum of 10%, 29% and 48% of the variances in the estimated serial intervals were 196 explained in the three waves respectively by both types of PHSMs (community-wide and case-197 based measures) when adjusting for the possible delays in the impact of onset-to-isolation. The 198 relative population mobility and daily tests per capita showed a positive and a negative 199 association, respectively, with the mean effective serial interval across the waves (figs. S7 to S9 and table S5). 200

201

## 202 Factor-specific effective serial intervals

203 We also explored the association between the estimated serial intervals and characteristics of 204 the infector in the identified transmission pairs in the second, third and fourth wave, including 205 age, sex, source of infection, transmission setting, severity outcomes and onset-to-isolation 206 delay (section 4, supplementary materials). The mean serial interval estimates ranged from 2.4 207 days to 5.6 days across the strata of these characteristic factors for infectors (tables S6 and S7, 208 and fig. S3). We found that an older age ( $\geq 65$  years) of the infectors was associated with a 209 higher estimate of the mean serial interval (p-value <0.023) across the waves (fig. S3A and 210 tables S7). The age-specific transmission matrices indicated that a higher proportion of the 211 secondary infections transmitted to younger (<35 years) infectees from the infectors of all age-212 groups (Fig. 2, A to C) and initially (during pre-peak) infections started with older (>65 years) 213 infector-infectee transmissions, then involved more younger population during peak and post-214 peak periods of the third and fourth waves (fig. S4). The age specific mean serial intervals were 215 longer for older infectors (e.g.,  $\geq 65$  years) (Fig 2, D to F and fig. S5).

216

217 In the second wave, transmission setting was identified as a statistically significant explanatory 218 factor (p-value = 0.037) with a higher estimated mean serial interval for household transmission 219 (tables S6 and S7). In general, the estimates of mean serial intervals were longer for the 220 transmission pairs with critical/severe infectors across the waves, and was statistically 221 significant for the third wave (p-value = 0.037) and fourth wave (p-value = 0.022). The onset-222 to-isolation intervals of the infectors were associated with estimated serial intervals (p-value < 223 0.005) across the waves, with shorter mean onset-to-isolation intervals correlating with shorter 224 mean serial interval estimates. Similar results were found in the sensitivity analysis for pre-225 defined thresholds from (5-14) days (table S6).

226

# 227 Correction in estimation of transmissibility accounting for time-varying effective serial 228 interval distributions

We estimated time-varying  $R_t$  to infer changes in the transmissibility of COVID-19 in Hong Kong using the Wallinga-Teunis method (4) (section 5, supplementary materials) which constructed the relative likelihood for a given case being a potential infector of other cases based on the case-based reproduction number (23). We extended the method by incorporating the effective serial interval distribution into the estimation of  $R_t$  via the *EpiEstim* package in R. The estimated  $R_t$  from time-varying effective serial intervals was compared with that based on a constant serial interval distribution over the epidemic waves.

236

237 A simulation based modelling framework was also applied to quantify the bias from using a constant over time-varying serial interval in estimating  $R_t$  through a comparison of the attack 238 239 rates (or cumulative number of infections) under both approaches (section 7, supplementary 240 materials). The time series of transmission rate  $\beta_t$  derived first from the  $R_t$ , estimated based on 241 time-varying effective serial interval distributions and possible choices of constant serial 242 interval distributions. Then these  $\beta_t$  were used to reconstruct the respective epidemic curves 243 by simulating Susceptible-Infected-Recovered models to assess the related bias as the 244 difference in attack rates from the reported data.

245

The differences between  $R_t$  estimated from real-time effective serial intervals and a constant serial interval distribution indicated that the latter might have introduced biases in inferring the transmissibility of COVID-19 over these 3 waves under analysis (**Fig. 3**). Smaller biases in  $R_t$ were identified during the fourth wave (except initial days) (**Fig. 3C**), during which effective serial interval distributions were relatively more stable (i.e., less varying) (**Fig. 1D**).

251 Furthermore, with possible choices of single constant mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of 252 serial interval distributions as estimated from our data, we predicted the bias in  $R_t$  by evaluating 253 the mean absolute deviation from the respective  $R_t$ , calculated by using effective serial 254 intervals. The magnitude and direction (over- or under-estimate) of these biases depends on the mean as well as the standard deviation of the fixed serial interval distributions (Fig. 3, A to C 255 256 and fig. S6). We found biases in simulated attack rates, were smaller when generated by 257 considering the effective serial interval distributions over the respective choices of constant 258 serial interval distributions (Fig. 3, D to F and table S8).

259

## 260 Discussion

261 The transmission dynamics and epidemiological characteristics of COVID-19 varied 262 significantly over time during the earlier epidemic waves studied in Hong Kong (Fig. 1 and 263 Fig. 3, A to C), accounting for changes in case profiles and the impact of various control 264 measures implemented at that time (table S1) (16, 17, 24). Our results for Hong Kong align 265 with earlier findings for mainland China, in identifying a shortening in mean serial intervals 266 over time as an epidemic is controlled with non-pharmaceutical interventions (9, 11). The first 267 two waves in Hong Kong were notably driven by imported cases and potential super-spreading 268 events (17), and the proportion of asymptomatic cases was found to be significantly higher 269 during later waves (15% - 23%). The proportion of pre-symptomatic transmissions was up to 270 10%, and was significantly higher during the third wave, and comparable with the reported 271 estimates elsewhere (18, 25). These inter- and intra-wave variations in case profiles along with 272 varied PHSMs could shape the epidemiological parameters and hence the transmission dynamics of COVID-19 waves in Hong Kong. 273

275 In early 2020, we reported substantial reductions in mean effective serial intervals over time in 276 mainland China following the implementation of various PHSMs including timely case 277 isolation (9). Here we observed similar results in the second wave in Hong Kong (Fig. 1B), 278 during which the mean onset-to-isolation delay had a positive association with mean serial 279 interval and could explain a high proportion (59%) of variance in effective serial intervals. 280 Other studies illustrated the effectiveness of the isolation of symptomatic cases and their 281 potential contacts in achieving sustained control of COVID-19 transmission (26, 27). Strict 282 isolation of cases will reduce transmission later in the infectious period (9-11). The measure of 283 case isolation improved (i.e., early isolation) over time in Hong Kong (28) and appeared to be 284 comparatively stable in third and fourth waves and have a lesser impact on the temporal 285 changes in mean serial intervals (tables S3 and S4), which might have been associated with the 286 broader application of PHSMs or demographic factors.

287

288 The community-wide PHSMs had negative association with serial interval across the waves, 289 while case-based PHSMs, including case isolation had positive association with serial interval 290 in the second and third wave (p-value = 0.33) (table S3). The association of these PHSMs with 291 effective serial intervals could be positive or negative based on the impact of these 292 interventions which could affect the infectors' infectiousness profile (as truncated or modified 293 by case-based measures including case isolation) and effective contact pattern of infectees with 294 potential infectors (as reduced by community-wide measures and mass testing) respectively (9-295 11). This finding could suggest the lengthening of mean serial intervals during post-peak of 296 third wave (during August 2020), when the per capita testing volume (as proxy of case-based 297 measures) started to decline (fig. S7 and table S5), which might be attributable to missing of 298 effective contact tracing and delays in isolation of cases (29, 30). The negative association 299 between serial interval and community-wide PHSMs (table S3), could delay in finding

infectees for an infector by reducing the probability of effective contacts and resulted in reverse
with the strengthening of community-wide PHSMs (*30, 31*).

302

303 Along with the PHSMs, we found that certain demographic characteristics of the infectors and 304 infectees were associated with variations in serial intervals (fig. S3A and tables S7). The older 305 cases might be more severe and had faster and higher viral load (32-39), which might have 306 been associated with shorter latent periods and incubation periods (36, 40-46). Therefore, the 307 serial intervals for the transmissions from older infectors to younger infectees could be the 308 longest or vice-versa as illustrated in fig. S10. For example, we noticed from early August 309 2021, during post-peak of third wave, the mean serial interval estimates started to be longer 310 although the onset-to-isolation intervals were more or less stable (Fig. 1C). Although we noticed on an average the infectors were older than the infectees in the transmission pairs 311 312 throughout the third wave and fourth waves (except during the end of fourth wave) (fig. S11). 313 Temporal variation in the age distributions of infector and infectee for a transmission chain 314 could reshape the serial interval (fig. S4). The mean infectors' age substantially increased after 315 the peak of third wave (from late June and early August 2020) and during the end of the fourth 316 wave (during February, 2021) (fig. S11), might lead longer mean serial intervals during these 317 periods (Fig. 1).

318

The infectors in household settings had significantly longer serial intervals for the third wave in Hong Kong (fig. S3 and table S7), accounting the proportion of older infectors in household was higher during that wave, though mean serial intervals under such settings reported not statistically significant for the first wave in mainland China (*9, 18*). For example, the household transmission was increased over time as noticed during third wave (fig. S12) and having olderto-younger transmission across the course of each wave (fig. S4) might lengthen the serial

intervals during post-peak of third wave, during which mean onset-to-isolation were muchstable (Fig. 1C).

327

328 Furthermore, the methods of estimation may affect the estimates of the serial interval 329 distribution (10). Estimating temporal serial intervals via forward-looking, backward-looking 330 or intrinsic approach by using a cohort-based framework have respective biases to be corrected 331 (10, 11, 47, 48). Therefore, the estimation of reproduction number with single constant serial 332 interval distribution can lead to bias (9) which can be minimised by using effective serial 333 interval distributions instead (Fig. 3, fig. S6 and table S8). This bias was noticeably lower in 334 fourth wave (Fig. 3) as by then the serial interval distribution had less variation across that 335 period (Fig. 1).

336

337 This study provides a unique opportunity for a temporal investigation of how mean serial 338 intervals can change over time, both shortening as well as lengthening depending on the 339 intensity of transmission and the various case-based and community-wide PHSMs applied in 340 Hong Kong, as well as variation in case profile (or characteristics), including demography, 341 clinical severity, and transmission settings between and across waves. This is the first study to 342 identify the potential factors, which could not only shorten the mean serial intervals over time, 343 but also could lengthen the measure, therefore the resulting temporal changes in the estimates 344 were accounted the direction and their strength of impacts. However, there are some limitations to our work. First, the change of the forward-looking serial intervals could be explained by 345 346 both the PHSMs and internal mechanism (backward incubation period of infectors) (11), but 347 detailed exposure information was not always available. Second, information on serial intervals 348 collected through contact tracing could suffer various biases such as recall biases, which could 349 reduce the chance of identifying longer serial intervals for example. However, our intense

algorithmic check and cross check with additional phylogenetic data should minimise the impact of this biases on the main outcomes. Finally, while we identified the potential for bias in  $R_t$  estimates if using constant serial interval distributions, there are other potential biases in  $R_t$  estimation that we did not discuss here.

354

In conclusion, our results indicated that the changes in serial interval distributions might not be 355 356 always monotonic as reported earlier for mainland China in 2020 (9). The real-time variations 357 in serial interval distributions were also driven by the changes in transmission pattern (who 358 infects whom), characterised by temporal variation in demographical and clinical profiles of 359 the cases along with PHSMs during the first two years of the COVID-19 pandemic in Hong 360 Kong. This time varying matric of effective serial interval distributions could improve the estimation of transmissibility, the time varying reproduction number  $(R_t)$  accounting for 361 362 intermediate factors of transmission and allow us to assess the timely impact of public health 363 measures.

364

365

#### **References and Notes** 367

369	1.	A. J. Kucharski et al., Early dynamics of transmission and control of COVID-19: a
370		mathematical modelling study. Lancet Infect Dis 20, 553-558 (2020).
371	2.	Y. Li et al., The temporal association of introducing and lifting non-pharmaceutical
372		interventions with the time-varying reproduction number (R) of SARS-CoV-2: a
373		modelling study across 131 countries. Lancet Infect Dis, (2020).
374	3.	A. Cori, N. M. Ferguson, C. Fraser, S. Cauchemez, A new framework and software to
375		estimate time-varying reproduction numbers during epidemics. Am J Epidemiol 178,
376		1505-1512 (2013).
377	4.	J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory
378		syndrome reveal similar impacts of control measures. Am J Epidemiol 160, 509-516
379		(2004).
380	5.	C. Fraser, Estimating individual and household reproduction numbers in an emerging
381		epidemic. <i>PLoS One</i> <b>2</b> , e758 (2007).
382	6.	A. Svensson, A note on generation times in epidemic models. <i>Math Biosci</i> <b>208</b> , 300-
383		311 (2007).
384	7.	M. Lipsitch <i>et al.</i> , Transmission dynamics and control of severe acute respiratory
385		syndrome. Science <b>300</b> , 1966-1970 (2003).
386	8.	B. J. Cowling, V. J. Fang, S. Riley, J. S. Malik Peiris, G. M. Leung, Estimation of the
387		serial interval of influenza. <i>Epidemiology</i> <b>20</b> , 344-347 (2009).
388	9.	S. T. Ali <i>et al.</i> , Serial interval of SARS-CoV-2 was shortened over time by
389		nonpharmaceutical interventions. Science <b>369</b> , 1106-1109 (2020).
390	10.	S. T. Ali <i>et al.</i> , Serial Intervals and Case Isolation Delays for Coronavirus Disease
391		2019: A Systematic Review and Meta-Analysis. Clin Infect Dis 74, 685-694 (2022).
392	11.	S. W. Park et al., Forward-looking serial intervals correctly link epidemic growth to
393		reproduction numbers. Proceedings of the National Academy of Sciences 118.
394		e2011548118 (2021).
395	12.	Centre for Health Protection (CHP) of the Department of Health, Hong Kong SAR,
396		Latest local situation of COVID-19, (31 December 2021)
397		https://www.chp.gov.hk/files/pdf/local situation covid19 en.pdf. (2021).
398	13.	B. Yang <i>et al.</i> , Changing Disparities in Coronavirus Disease 2019 (COVID-19)
399		Burden in the Ethnically Homogeneous Population of Hong Kong Through Pandemic
400		Waves: An Observational Study. Clin Infect Dis 73, 2298-2305 (2021).
401	14.	Y. M. Mefsin <i>et al.</i> , Epidemiology of Infections with SARS-CoV-2 Omicron BA.2
402		Variant, Hong Kong, January-March 2022. Emerg Infect Dis 28, (2022).
403	15.	Hong Kong Government News for COVID-19 (31 December, 2021)
404		https://www.news.gov.hk/eng/categories/covid19/index.html). (2021).
405	16.	P. Wu <i>et al.</i> , Suppressing COVID-19 transmission in Hong Kong: an observational
406		study of the first four months. 09 June 2020, PREPRINT (Version 1) available at
407		SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3627304, (2020).
408	17.	D. C. Adam et al., Clustering and superspreading potential of SARS-CoV-2 infections
409		in Hong Kong. Nat Med 26, 1714-1719 (2020).
410	18.	X. K. Xu et al., Reconstruction of Transmission Pairs for novel Coronavirus Disease
411		2019 (COVID-19) in mainland China: Estimation of Super-spreading Events, Serial
412		Interval, and Hazard of Infection. Clin Infect Dis, (2020).
413	19.	D. Klinkenberg, J. A. Backer, X. Didelot, C. Colijn, J. Wallinga, Simultaneous
414		inference of phylogenetic and transmission trees in infectious disease outbreaks. PLoS
415		<i>Comput Biol</i> <b>13</b> , e1005495 (2017).

416 417	20.	K. Leung, J. T. Wu, G. M. Leung, Real-time tracking and prediction of COVID-19 infection using digital proxies of population mobility and mixing. <i>Nat Commun</i> <b>12</b> , 1501 (2021)
418	0.1	1501 (2021).
419	21.	M. U. G. Kraemer <i>et al.</i> , The effect of human mobility and control measures on the
420	~~	COVID-19 epidemic in China. <i>Science</i> <b>368</b> , 493-497 (2020).
421	22.	C. O. Buckee <i>et al.</i> , Aggregated mobility data could help fight COVID-19. <i>Science</i>
422		<b>368</b> , 145-146 (2020).
423	23.	K. M. Gostic <i>et al.</i> , Practical considerations for measuring the effective reproductive
424		number, Rt. <i>PLoS Comput Biol</i> <b>16</b> , e1008409 (2020).
425	24.	B. J. Cowling <i>et al.</i> , Impact assessment of non-pharmaceutical interventions against
426		coronavirus disease 2019 and influenza in Hong Kong: an observational study. <i>Lancet</i>
427		Public Health 5, e279-e288 (2020).
428	25.	Z. Du <i>et al.</i> , Serial Interval of COVID-19 among Publicly Reported Confirmed Cases.
429		<i>Emerg Infect Dis</i> <b>26</b> , 1341-1343 (2020).
430	26.	Y. Ng et al., Evaluation of the Effectiveness of Surveillance and Containment
431		Measures for the First 100 Patients with COVID-19 in Singapore - January 2-
432		February 29, 2020. MMWR Morb Mortal Wkly Rep 69, 307-311 (2020).
433	27.	A. J. Kucharski <i>et al.</i> , Effectiveness of isolation, testing, contact tracing, and physical
434		distancing on reducing transmission of SARS-CoV-2 in different settings: a
435		mathematical modelling study. <i>Lancet Infect Dis</i> <b>20</b> , 1151-1160 (2020).
436	28.	H. Y. Lam <i>et al.</i> , The epidemiology of COVID-19 cases and the successful
437		containment strategy in Hong Kong-January to May 2020. Int J Infect Dis 98, 51-58
438		(2020).
439	29.	B. Yang <i>et al.</i> , Universal community nucleic acid testing for COVID-19 in Hong
440		Kong reveals insights into transmission dynamics: a cross-sectional and modelling
441		study. Clin Infect Dis, (2021).
442	30.	H. Y. Yuan, C. Blakemore, The impact of multiple non-pharmaceutical interventions
443		on controlling COVID-19 outbreak without lockdown in Hong Kong: A modelling
444		study. Lancet Reg Health West Pac 20, 100343 (2022).
445	31.	K. Leung, J. T. Wu, G. M. Leung, Effects of adjusting public health, travel, and social
446		measures during the roll-out of COVID-19 vaccination: a modelling study. <i>Lancet</i>
447		<i>Public Health</i> <b>6</b> , e674-e682 (2021).
448	32.	F. Shi <i>et al.</i> , Association of viral load with serum biomakers among COVID-19 cases.
449		<i>Virology</i> <b>546</b> , 122-126 (2020).
450	33.	F. X. Lescure <i>et al.</i> , Clinical and virological data of the first cases of COVID-19 in
451	- ·	Europe: a case series. Lancet Infect Dis 20, 697-706 (2020).
452	34.	Y. Liu <i>et al.</i> , Viral dynamics in mild and severe cases of COVID-19. <i>Lancet Infect</i>
453	~ -	Dis 20, 656-657 (2020).
454	35.	Y. Pan, D. Zhang, P. Yang, L. L. M. Poon, Q. Wang, Viral load of SARS-CoV-2 in
455		clinical samples. Lancet Infect Dis 20, 411-412 (2020).
456	36.	K. A. Walsh <i>et al.</i> , SARS-CoV-2 detection, viral load and infectivity over the course
457		of an infection. J Infect 81, 357-371 (2020).
458	37.	Y. D. Gao <i>et al.</i> , Risk factors for severe and critically ill COVID-19 patients: A
459	•	review. Allergy <b>76</b> , 428-455 (2021).
460	38.	K. K. To <i>et al.</i> , Temporal profiles of viral load in posterior oropharyngeal saliva
461		samples and serum antibody responses during infection by SARS-CoV-2: an
462	•	observational cohort study. Lancet Infect Dis 20, 565-574 (2020).
463	39.	E. Pujadas <i>et al.</i> , SARS-CoV-2 viral load predicts COVID-19 mortality. <i>Lancet</i>
464		<i>Respir Med</i> <b>8</b> , e70 (2020).

465	5 40.	T. Ward, A. Johnsen, Understanding an evolving pandemic: An analysis of the
460	)	clinical time delay distributions of COVID-19 in the United Kingdom. PLoS One 16,
46/	/	e0257978 (2021).
468	3 41.	M. Levine-Tiefenbrun <i>et al.</i> , Initial report of decreased SARS-CoV-2 viral load after
469	)	inoculation with the BNT162b2 vaccine. <i>Nat Med</i> 27, 790-792 (2021).
470	) 42.	H. Xin <i>et al.</i> , The incubation period distribution of coronavirus disease 2019
471	[	(COVID-19): a systematic review and meta-analysis. <i>Clin Infect Dis</i> , (2021).
472	2 43.	C. Leung, The difference in the incubation period of 2019 novel coronavirus (SARS-
473	3	CoV-2) infection between travelers to Hubei and nontravelers: The need for a longer
474	1	quarantine period. Infect Control Hosp Epidemiol 41, 594-596 (2020).
475	5 44.	X. He <i>et al.</i> , Temporal dynamics in viral shedding and transmissibility of COVID-19.
476	5	Nat Med 26, 672-675 (2020).
477	7 45.	M. Levine-Tiefenbrun et al., SARS-CoV-2 RT-qPCR Test Detection Rates Are
478	3	Associated with Patient Age, Sex, and Time since Diagnosis. J Mol Diagn, (2021).
479	9 46.	T. C. Jones <i>et al.</i> , Estimating infectiousness throughout SARS-CoV-2 infection
480	)	course. Science 373, (2021).
481	l 47.	D. Champredon, J. Dushoff, Intrinsic and realized generation intervals in infectious-
482	2	disease transmission. Proc Biol Sci 282, 20152026 (2015).
483	3 48.	S. W. Park, D. Champredon, J. S. Weitz, J. Dushoff, A practical generation-interval-
484	1	based approach to inferring the strength of epidemics from their speed. <i>Epidemics</i> 27.
485	5	12-18 (2019).
486	6 49.	X. F. Liu, X. K. Xu, Y. Wu, Mobility, exposure, and epidemiological timelines of
487	7	COVID-19 infections in China outside Hubei province. Sci Data 8, 54 (2021).
488	<b>5</b> 0.	S. T. Ali <i>et al.</i> , Serial intervals and case isolation delays for COVID-19: a systematic
489	)	review and meta-analysis. Clin Infect Dis, ciab491 (2021).
490	) 51.	J. T. Wu et al., Nowcasting epidemics of novel pathogens: lessons from COVID-19.
491	l	Nat Med 27, 388-395 (2021).
492	2 52.	R. Subramanian, Q. He, M. Pascual, Quantifying asymptomatic infection and
493	3	transmission of COVID-19 in New York City using observed cases, serology, and
494	1	testing capacity. Proc Natl Acad Sci US A 118, (2021).
495	5 53.	S. T. Ali, B. J. Cowling, E. H. Y. Lau, V. J. Fang, G. M. Leung, Mitigation of
496	6	Influenza B Epidemic with School Closures, Hong Kong, 2018. Emerg Infect Dis 24,
497	7	2071-2073 (2018).
498	3 54.	S. T. Ali <i>et al.</i> , Prediction of Upcoming Global Infection Burden of Influenza Seasons
499	)	after Relaxation of Public Health and Social Measures for COVID-19 Pandemic.
500	)	Available at SSRN: https://ssrn.com/abstract=4063811 or
501	[	http://dx.doi.org/10.2139/ssrn.4063811, (2022).
502	2 55.	L. Sigfrid <i>et al.</i> , What is the recovery rate and risk of long-term consequences
503	3	following a diagnosis of COVID-19? A harmonised global longitudinal observational
504	1	study protocol. <i>BMI Open</i> 11, e043887 (2021)
505	5	
202	,	

## 508 Acknowledgments:

We acknowledge Ms. Julie Au for technical assistance. We thank the Centre for Health Protection in Hong Kong for their data collection and compilation throughout the COVID-19 pandemic. We thank Kathy Leung for helpful discussions. STA, DC, and acknowledge the research computing facilities and advisory services offered by Information Technology Services, The University of Hong Kong.

514

## 515 **Funding:**

- 516 Health and Medical Research Fund, Health Bureau, Government of the Hong Kong Special
- 517 Administrative Region grant COVID190118 (BJC) and grant 20190712 (STA);
- 518 Collaborative Research Scheme grant C7123-20G and grant T11-705/14N, Research Grants
- 519 Council, Government of the Hong Kong Special Administrative Region (BJC);
- 520 AIR@innoHK program of the Innovation and Technology Commission, Government of the
- 521 Hong Kong Special Administrative Region (STA, ZD, EHYL, PW, GML, BJC);
- 522 European Research Council (grant no. 804744); the Grand Challenges ICODA pilot initiative,
- 523 delivered by Health Data Research UK and funded by the Bill & Melinda Gates Foundation
- and the Minderoo Foundation (LW);
- 525 The funding bodies had no role in study design, data collection and analysis, preparation of the 526 manuscript, or the decision to publish.
- 527

## 528 Author contributions:

- 529 Conceptualization: STA, PW, BJC
- 530 Methodology: STA, DC, LW, EHYL, BJC
- 531 Investigation: STA, DC, WWL, AY, DCA, YCL
- 532 Data: WWL, AY, DC, DCA, JX, FH, HG, STA
- 533 Visualization: DC, AY, YCL, STA
- 534 Comments: DCA, EHYL, LW, ZD, JYW, XKX, PW, GML, BJC
- 535 Funding acquisition: STA, EHYL, ZD, LW, PW, GML, BJC
- 536 Project administration: STA, DC, WWL, DCA, YCL
- 537 Supervision: GML, BJC

- 538 Writing original draft: STA, DC, BJC
- 539 Writing review & editing: STA, DCA, EHYL, LW, ZD, JYW, XKX, PW, GML, BJC
- 540

# 541 **Competing interests:**

- 542 BJC consults for AstraZeneca, Fosun Pharma, GSK, Moderna, Pfizer, Roche and Sanofi
- 543 Pasteur. The authors report no other potential conflicts of interest.
- 544

# 545 Data and materials availability:

- 546 All data, code, and materials used in the analyses in main text and supplementary materials
- 547 will be available together with the publication of this paper.

548

# 549 Supplementary Materials

- 550 Materials and Methods
- 551 Supplementary Text
- 552 Figs. S1 to S12
- 553 Tables S1 to S8
- 554 References (49–55)

555

556

557



Fig. 1. Transmission dynamics and temporal estimates of serial interval distribution and 561 562 isolation delay for COVID-19 in Hong Kong. (A) The grey bars indicate the epi-curve of the reported COVID-19 cases, and white bars are onset epi-curves for infectors of those reported 563 564 cases during four waves in Hong Kong. Mean serial interval estimates (red dots) with 95% CrI 565 (red vertical line segments) and mean onset-to-isolation interval (teal dots) with 95% CrI (teal 566 vertical line segments), evaluated from all confirmed and likely transmission pairs during pre-567 peak, peak-timing, and post-peaks for three pre-defined waves (light blue shades) with vertical 568 dark grey dashed lines referring to the peak-timing for each wave, with two peaks for fourth 569 wave. (B)-(D) Time varying estimates of effective serial intervals and onset-to-isolation 570 interval for second wave (A), third wave (B) and fourth wave (C) with the indicator to timings 571 of major public health and social measures (PHSMs) implemented in Hong Kong. The area 572 between two grey dashed lines in each wave indicates the peak timing of the epidemic wave. 573 The estimates of serial interval and onset-to-isolation interval are evaluated by using MCMC 574 on fitting normal and gamma distributions to empirical data on confirmed and likely 575 transmission pairs (considering predefined threshold of 8 days) respectively. The effective 576 serial intervals and onset-to-isolation intervals are estimated based on the empirical data in 10day sliding windows (presented with respect to the 5<sup>th</sup> day of each sliding window). 577

578

579

580



Fig. 2. Age-specific transmission and serial interval estimated between infectors and 583 584 infectees for COVID-19 in Hong Kong. The heat maps for age-specific transmission densities (the relative frequency matrices of the age distribution of infector-infectee transmission pairs 585 for the age-groups of below 35, 35-45, 45-55, 55-65, above 65 years) with the marginal 586 587 densities were for third wave (A), 1<sup>st</sup> peak of fourth wave (B) and 2<sup>nd</sup> peak of fourth wave (C) of COVID-19 in Hong Kong. Age-specific empirical mean serial intervals with the respective 588 589 marginal estimates, evaluated for these age-groups stratifications for infector and infectee across the third wave (D), 1<sup>st</sup> peak of fourth waves (E) and 2<sup>nd</sup> peak of fourth wave (F). Note 590 that no infector found of age 65 years and over during second wave, hence excluded for age-591 592 specific analysis.



604 605 Fig. 3. Correction in transmissibility for the bias from using constant serial interval over 606 effective serial interval distributions. Comparison of daily estimates of effective reproduction numbers  $(R_t)$  by using real-time effective serial interval distributions versus using 607 a single fixed serial interval distributions for second wave (A), third wave (B) and fourth wave 608 609 (C) in Hong Kong. The black solid lines are mean  $R_t$  (with 95% CI in dashed black lines), 610 evaluated using effective serial interval distributions and other solid colours lines represent the 611 mean  $R_t$  (with 95% CI in respective colours shades), evaluated using single fixed serial interval 612 distributions. The fixed serial interval distributions with means (standard deviations) were 613 considered for pre-, during and post-peak respectively in each wave as mean 5.5 (2.4) days, 4.5 (3.2) days and 3.2 (2.5) days for second wave; 4.6 (3.9) days, 3.1 (3.2) days and 2.7 (4.3) days 614 615 for third wave; 4.0 (2.8) days, 3.5 (3.1) days and 4.0 (2.4) days for 1<sup>st</sup> peak of fourth wave; and 4.0 (3.0) days, 3.1 (2.7) days and 3.2 (3.3) days for 2<sup>nd</sup> peak of fourth wave. The differences 616 617 between the black solid lines and the coloured solid lines indicate the respective biases on using 618 single constant serial interval distribution to estimate  $R_t$ . (D)-(E) The biases in terms of attack 619 rates, estimated by reconstructing the epi-curves via. using the transmissibility  $(R_t)$  derived by 620 effective and constants serial interval distributions and Susceptible-Infected-Recovered (SIR) 621 models for second wave (A), third wave (B) and fourth wave (C) in Hong Kong. The observed 622 epi-curves (in grey bars) of the COVID-19 cases, the reconstructed epi-curves in dashed lines by using effective serial interval distributions (in black dashed lines) and different 623 624 counterfactual constant serial interval distributions (in dashed coloured lines). The differences 625 between the black dashed lines and the coloured dashed lines indicate the respective biases in 626 attack rates on using constant serial interval distributions. 627