Heterogeneity in transmissibility and shedding SARS-CoV-2 via droplets and aerosols

Paul Z. Chen[a], Niklas Bobrovitz[b-d], Zahra Premji[e], Marion Koopmans[f], David N. Fisman[g,h], Frank X. Gu[a,i,1]

Affiliations:

[a]Department of Chemical Engineering & Applied Chemistry, University of Toronto, Toronto, Canada

[b]Temerty Faculty of Medicine, University of Toronto, Toronto, Canada

[c]Department of Critical Care Medicine, Cumming School of Medicine, University of Calgary, Calgary, Canada

[d]O'Brien Institute of Public Health, University of Calgary, Calgary, Canada

[e]Libraries & Cultural Resources, University of Calgary, Calgary, Canada

[f]Department of Viroscience, Erasmus University Medical Center, Rotterdam, Netherlands

[g]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[h]Division of Infectious Diseases, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada

[i]Institute of Biomedical Engineering, University of Toronto, Toronto, Canada

[1]To whom correspondence may be addressed: Frank X. Gu.

Email: f.gu@utoronto.ca

Abstract

Which virological factors mediate overdispersion in the transmissibility of emerging viruses remains a longstanding question in infectious disease epidemiology. Here, we use systematic review to develop a comprehensive dataset of respiratory viral loads (rVLs) of SARS-CoV-2, SARS-CoV-1 and influenza A(H1N1)pdm09. We then comparatively meta-analyze the data and model individual infectiousness by shedding viable virus via respiratory droplets and aerosols. Our analyses indicate heterogeneity in rVL as an intrinsic virological factor facilitating greater overdispersion for SARS-CoV-2 in the COVID-19 pandemic than A(H1N1)pdm09 in the 2009 influenza pandemic. For COVID-19, case heterogeneity remains broad throughout the infectious period, including for pediatric and asymptomatic infections. Hence, many COVID-19 cases inherently present minimal transmission risk, whereas highly infectious individuals shed tens to thousands of SARS-CoV-2 virions/min via droplets and aerosols while breathing, talking and singing. Coughing increases the contagiousness, especially in close contact, of symptomatic cases relative to asymptomatic ones. Infectiousness tends to be elevated between 1-5 days post-symptom onset. Our findings show how individual case variations influence virus transmissibility and present considerations for disease control in the COVID-19 pandemic.

Significance Statement

For some emerging infectious diseases, including COVID-19, few cases cause most secondary infections. Others, like influenza A(H1N1)pdm09, spread more homogenously. The virological factors that mediate such distinctions in transmissibility remain unelucidated, prohibiting the development of specific disease control measures. We find that intrinsic case variation in respiratory viral load (rVL) facilitates overdispersion, and superspreading, for COVID-19 but

more homogeneous transmission for A(H1N1)pdm09. We interpret the influence of heterogeneity in rVL on individual infectiousness by modelling likelihoods of shedding viable virus via respiratory droplets and aerosols. We analyze the distribution and kinetics of SARS-CoV-2 rVL, including across age and symptomatology subgroups. Our findings compare individual infectiousness across COVID-19 and A(H1N1)pdm09 cases and present quantitative guidance on triaging COVID-19 contact tracing.

Main Text

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread globally, causing the coronavirus disease 2019 (COVID-19) pandemic with more than 62.5 million infections and 1.4 million deaths (as of 29 November 2020) (1). For respiratory virus transmission, airway epithelial cells shed virions to the extracellular fluid before atomization (from breathing, talking, singing, coughing and aerosol-generating procedures) partitions them into a polydisperse mixture of particles that are expelled to the ambient environment. Aerosols (≤100 μm) can be inhaled nasally, whereas droplets (>100 μm) tend to be excluded (2, 3). For direct transmission, droplets must be sprayed ballistically onto susceptible tissue (4). Hence, droplets predominantly deposit on nearby surfaces, potentiating indirect transmission. Aerosols can be further categorized based on typical travel characteristics: short-range aerosols (50-100 μm) tend to settle within 2 m; long-range ones (10-50 μm) often travel beyond 2 m based on emission force; and buoyant aerosols (≤10 μm) remain suspended and travel based on airflow profiles for minutes to many hours (4, 5). Although proximity has been associated with infection risk for COVID-19 (6), studies have also suggested that long-range airborne transmission occurs conditionally (7-9).

68 While the basic reproductive number has been estimated to be 2.0-3.6 (10, 11),
69 transmissibility of SARS-CoV-2 is highly overdispersed (dispersion parameter *k*, 0.10-0.58),
70 with numerous instances of superspreading (7-9) and few cases (10-20%) causing many
71 secondary infections (80%) (12-14). Similarly, few cases drive the transmission of SARS-CoV-1
72 (*k*, 0.16-0.17) (15), whereas influenza A(H1N1)pdm09 transmits more homogeneously (*k*, 7.4-
73 14.4) (16, 17), despite both viruses spreading by contact, droplets and aerosols (18, 19).
74 Although understanding the determinants of viral overdispersion is crucial towards
75 characterizing transmissibility and developing effective strategies to limit infection (20),
76 mechanistic associations for *k* remain unclear. As an empirical estimate, *k* depends on myriad
77 extrinsic (behavioral, environmental and invention) and host factors. Nonetheless, *k* remains
78 similar across distinct outbreaks (15), suggesting that intrinsic virological factors mediate
79 overdispersion for emerging viruses.
80 Here, we investigated how intrinsic case variation in rVL facilitates overdispersion in
81 SARS-CoV-2 transmissibility. By systematic review, we developed a comprehensive dataset of
82 rVLs from cases of COVID-19, SARS and A(H1N1)pdm09. Using comparative meta-analyses,
83 we found that heterogeneity in rVL was associated with overdispersion among these emerging
84 infections. To assess potential sources of case heterogeneity, we analyzed SARS-CoV-2 rVLs
85 across age and symptomatology subgroups as well as disease course. To interpret the influence
86 of heterogeneity in rVL on individual infectiousness, we modelled likelihoods of shedding viable
87 virus via respiratory droplets and aerosols.
88
89 **Results**
90 **Systematic review**

91 We conducted a systematic review on virus quantitation in respiratory specimens taken during
92 the infectious periods of SARS-CoV-2 (-3 to 10 days from symptom onset [DFSO]) (21-23),
93 SARS-CoV-1 (0-20 DFSO) (24) and A(H1N1)pdm09 (-2 to 9 DFSO) (25) (Methods). The
94 systematic search (*SI Appendix*, Tables S1 to S5) identified 4,274 results. After screening and
95 full-text review, 64 studies met the inclusion criteria and contributed to the systematic dataset
96 (Fig. 1) (N = 9,631 total specimens), which included adult (N = 5,033) and pediatric (N = 1,608)
97 cases from 15 countries and specimen measurements for asymptomatic (N = 2,387),
98 presymptomatic (N = 28) and symptomatic (N = 7,161) infections. According to a hybrid Joanna
99 Briggs Institute critical appraisal checklist, risk of bias was low for most contributing studies (*SI*
100 *Appendix*, Table S6).
101
102 **Association of overdispersion with heterogeneity in rVL**
103 We hypothesized that individual case variation in rVL facilitates the distinctions in k among
104 COVID-19, SARS and A(H1N1)pdm09. For each study in the systematic dataset, we used
105 specimen measurements (viral RNA concentration in a respiratory specimen) to estimate rVLs
106 (viral RNA concentration in the respiratory tract) (Methods). To investigate the relationship
107 between k and heterogeneity in rVL, we performed a meta-regression using each contributing
108 study (*SI Appendix*, Fig. S1), which showed a weak, negative association between the two
109 variables (meta-regression slope t-test: P = 0.038, Pearson's r = -0.26).
110 Using contributing studies with low risk of bias, meta-regression (Fig. 2) showed a strong,
111 negative association between k and heterogeneity in rVL (meta-regression slope t-test: $P < 10^{-9}$,
112 Pearson's r = -0.73). In this case, each unit increase (1 $\log_{10}$ copies/ml) in the standard deviation
113 (SD) of rVL decreased log(k) by a factor of -1.41 (95% confidence interval [CI]: -1.78 to -1.03),

114 suggesting that broader heterogeneity in rVL facilitates greater overdispersion in the
115 transmissibility of SARS-CoV-2 than of A(H1N1)pmd09. To investigate mechanistic aspects of
116 this association, we conducted a series of analyses on rVL and then modelled the influence of
117 heterogeneity in rVL on individual infectiousness.
118

119 **Meta-analysis and subgroup analyses of rVL**

120 We first compared rVLs among the emerging infections. We performed a random-effects meta-
121 analysis (*SI Appendix*, Fig. S2), which assessed the expected rVL when encountering a COVID-
122 19, SARS or A(H1N1)pdm09 case during the infectious period. This showed that the expected
123 rVL of SARS-CoV-2 was comparable to that of SARS-CoV-1 (one-sided Welch's *t*-test: P =
124 0.111) but lesser than that of A(H1N1)pdm09 (P = 0.040). We also performed random-effects
125 subgroup analyses for COVID-19 (Fig. 3), which showed that expected SARS-CoV-2 rVLs were
126 similar between pediatric and adult cases (P = 0.476) and comparable between
127 symptomatic/presymptomatic and asymptomatic infections (P = 0.090). Since these meta-
128 analyses had significant between-study heterogeneity among the mean estimates (Cochran's Q
129 test: $P < 0.001$ for each meta-analysis), we conducted risk-of-bias sensitivity analyses; meta-
130 analyses of low-risk-of-bias studies continued to show significant heterogeneity (*SI Appendix*,
131 Figs. S3 to S7).
132

133 **Distributions of rVL**

134 We next analyzed rVL distributions. For all three viruses, rVLs best conformed to Weibull
135 distributions (*SI Appendix*, Fig. S8), and we fitted the entirety of individual sample data for each
136 virus in the systematic dataset (Fig. 4*A* and *SI Appendix*, Fig. S8*N*). While COVID-19 and SARS

137 cases tended to shed lesser virus than those with A(H1N1)pdm09 (*SI Appendix*, Fig. S2), broad
138 heterogeneity in SARS-CoV-2 and SARS-CoV-1 rVLs inverted this relationship for highly
139 infectious individuals (Fig. 4*A* and *SI Appendix*, Fig. S9 *A* to *C*). At the 90th case percentile (cp)
140 throughout the infectious period, the estimated rVL was 8.91 (95% CI: 8.83-9.00) $\log_{10}$
141 copies/ml for SARS-CoV-2, whereas it was 8.62 (8.47-8.76) $\log_{10}$ copies/ml for A(H1N1)pdm09
142 (*SI Appendix*, Table S7). The SD of the overall rVL distribution for SARS-CoV-2 was 2.04 $\log_{10}$
143 copies/ml, while it was 1.45 $\log_{10}$ copies/ml for A(H1N1)pdm09, showing that heterogeneity in
144 rVL was indeed broader for SARS-CoV-2.
145 To assess potential sources of heterogeneity in SARS-CoV-2 rVL, we compared rVL
146 distributions among COVID-19 subgroups. In addition to comparable mean estimates (Fig. 3),
147 adult, pediatric, symptomatic/presymptomatic and asymptomatic COVID-19 cases showed
148 similar rVL distributions (Fig. 4 *B* and *C*), with SDs of 2.03, 2.06, 2.00 and 2.01 $\log_{10}$ copies/ml,
149 respectively. Thus, age and symptomatology minimally influenced case variation in SARS-CoV-
150 2 rVL during the infectious period.
151
152 **SARS-CoV-2 kinetics during respiratory infection**
153 To analyze the influences of disease course, we delineated individual SARS-CoV-2 rVLs by
154 DFSO and fitted the mean estimates to a mechanistic model for respiratory virus kinetics (Fig.
155 4*D* and Methods). The outputs indicated that, on average, each productively infected cell in the
156 airway epithelium shed SARS-CoV-2 at 1.33 (95% CI: 0.74-1.93) copies/ml day^{-1} and infected
157 up to 9.25 susceptible cells (*SI Appendix*, Table S8). The turnover rate for infected epithelial
158 cells was 0.71 (0.26-1.15) days^{-1}, while the half-life of SARS-CoV-2 in the respiratory tract was
159 5.04 (2.62-66.0) hours. By extrapolating the model to an initial rVL of 0 $\log_{10}$ copies/ml, the

160 estimated incubation period was 5.38 days, which agrees with epidemiological findings (10).
161 Conversely, the expected duration of shedding was 25.1 DFSO. Thus, SARS-CoV-2 rVL
162 increased exponentially after infection, peaked around 1 DFSO along with the proportion of
163 infected epithelial cells (*SI Appendix*, Fig. S10) and then diminished exponentially.
164 To evaluate case heterogeneity across the infectious period, we fitted distributions for each
165 DFSO (Fig. 4*E*), which showed that high SARS-CoV-2 rVLs also increased from the
166 presymptomatic period, peaked at 1 DFSO and then decreased towards the end of the first week
167 of illness. For the 90th cp at 1 DFSO, the rVL was 9.84 (95% CI: 9.17-10.56) $\log_{10}$ copies/ml, an
168 order of magnitude greater than the overall 90th-cp estimate. High rVLs between 1-5 DFSO were
169 elevated above the expected values from the overall rVL distribution (*SI Appendix*, Table S7). At
170 -1 DFSO, the 90th-cp rVL was 8.30 (6.88-10.02) $\log_{10}$ copies/ml, while it was 7.93 (7.35-8.56)
171 $\log_{10}$ copies/ml at 10 DFSO. Moreover, heterogeneity in rVL remained broad across the
172 infectious period, with SDs of 1.83-2.44 $\log_{10}$ copies/ml between -1 to 10 DFSO (*SI Appendix*,
173 Fig. S9 *H* to *S*).
174
175 **Likelihood that droplets and aerosols contain virions**
176 Towards analyzing the influence of heterogeneity in rVL on individual infectiousness, we first
177 modelled the likelihood of respiratory particles containing viable SARS-CoV-2. Since rVL is an
178 intensive quantity, the volume fraction of virions is low and viral partitioning coincides with
179 atomization, we used Poisson statistics to model likelihood profiles. To calculate an unbiased
180 estimator of partitioning (the expected number of viable copies per particle), our method
181 multiplied rVL estimates with particle volumes during atomization and an assumed viability
182 proportion of 0.1% in equilibrated particles (Methods).

183 When expelled by the mean COVID-19 case during the infectious period, respiratory
184 particles showed low likelihoods of carrying viable SARS-CoV-2 (*SI Appendix*, Fig. S11*B*).
185 Aerosols (equilibrium aerodynamic diameter [d_a] ≤ 100 μm) were ≤0.69% (95% CI: 0.43-0.95%)
186 likely to contain a virion. Droplets also had low likelihoods: at d_a = 330 μm, they were 19.4%
187 (18.5-20.3%), 2.42% (2.15-2.69%) and 0.20% (0.16-0.24%) likely to contain one, two or three
188 virions, respectively.
189 COVID-19 cases with high rVLs, however, expelled particles with considerably greater
190 likelihoods of carrying viable copies (Fig. 5 *A* and *B* and *SI Appendix*, Fig. S11 *D* and *E*). For the
191 80th cp during the infectious period, aerosols (d_a ≤ 100 μm) were ≤36.6% (95% CI: 35.3-38.0%)
192 likely to carry at least one SARS-CoV-2 virion. For the 90th cp, this likelihood was ≤96.7%
193 (96.5-96.9%), with larger aerosols tending to contain multiple virions (*SI Appendix*, Fig. S11*E*).
194 At 1 DFSO, these estimates were greatest, and ≤61.1% (51.8-70.4%) of buoyant aerosols (d_a ≤
195 10 μm) contained at least one viable copy of SARS-CoV-2 for the 98th cp. When expelled by
196 high cps, droplets (d_a > 100 μm) tended to contain tens to thousands of SARS-CoV-2 virions
197 (Fig. 5*B* and *SI Appendix*, Fig. S11*E*).
198
199 **Shedding SARS-CoV-2 via respiratory droplets and aerosols**
200 Using the partitioning estimates in conjunction with published profiles of the particles expelled
201 by respiratory activities (*SI Appendix*, Fig. S12), we next modelled the rates at which talking,
202 singing, breathing and coughing shed viable SARS-CoV-2 across d_a (Fig. 5 *C* to *F*). Singing
203 shed virions more rapidly than talking based on the increased emission of aerosols. Voice
204 amplitude, however, had a significant effect on aerosol production, and talking loudly emitted
205 aerosols at similar rates to singing (*SI Appendix*, Fig. S12*E*). Based on the generation of larger

206 aerosols and droplets, talking and singing shed virions more rapidly than breathing (Fig. 5 *C* to
207 *E*). Each cough shed similar quantities of virions as in a minute of talking (Fig. 5 *C* and *F*).
208 Each of these respiratory activities expelled aerosols at greater rates than droplets, but
209 particle size correlated with the likelihood of containing virions. Talking, singing and coughing
210 expelled virions at greater proportions via droplets (80.6-86.0%) than aerosols (14.0-19.4%)
211 (Fig. 5*G*). Moreover, short-range aerosols predominantly mediated the virions (90.8-92.6%) shed
212 via aerosols while talking normally and coughing. In comparison, while singing, talking loudly
213 and breathing, buoyant (29.7-68.4%) and long-range (10.3-31.6%) aerosols carried a larger
214 proportion of the virions shed via aerosols (Fig. 5*G*).
215
216 **Influence of heterogeneity in rVL on individual infectiousness**
217 To interpret how heterogeneity in rVL influences individual infectiousness, we modelled total
218 SARS-CoV-2 shedding rates (over all particle sizes) for each respiratory activity (Fig. 5*H* and *SI*
219 *Appendix*, Fig. S13). Between the 1st and the 99th cps, the estimates for a respiratory activity
220 spanned ≥8.48 orders of magnitude on each DFSO; cumulatively from -1 to 10 DFSO, they
221 spanned 11.0 orders of magnitude. Hence, many COVID-19 cases inherently presented minimal
222 transmission risk, whereas highly infectious individuals shed considerable quantities of SARS-
223 CoV-2. For the 98th cp at 1 DFSO, singing expelled 313 (95% CI: 37.5-3,158) virions/min to the
224 ambient environment, talking emitted 293 (35.1-2,664) virions/min, breathing exhaled 1.54
225 (0.18-15.5) virions/min and coughing discharged 249 (29.8-2,5111) virions/cough; these
226 estimates were approximately two orders of magnitude greater than those for the 85th cp. For the
227 98th cp at -1 DFSO, singing shed 14.5 (0.15-4,515) virions/min and breathing exhaled 7.13×10^{-2}
228 (7.20×10^{-4}-22.2) virions/min. The estimates at 9-10 DFSO were similar to these presymptomatic

229 ones (Fig. 5*H* and *SI Appendix*, Fig. S13*B*). As indicated by comparable mean rVLs (Fig. 3) and
230 heterogeneities in rVL (Fig. 4 *B* and *C*), adult, pediatric, symptomatic/presymptomatic and
231 asymptomatic COVID-19 subgroups presented similar distributions for shedding virions through
232 these activities.

233 We also compared the influence of case variation on individual infectiousness between
234 A(H1N1)pdm09 and COVID-19. Aerosol spread accounted for approximately half of
235 A(H1N1)pdm09 transmission events (19), and the 50% human infectious dose for aerosolized
236 influenza A virus is approximately 1-3 virions in the absence of neutralizing antibodies (26).
237 Based on the model, 62.9% of A(H1N1)pdm09 cases were infectious (shed ≥1 virion) via
238 aerosols within 24 h of talking loudly or singing (*SI Appendix*, Fig. S14*A*). The estimate was
239 58.6% within 24 h of talking normally and 22.3% within 24 h of breathing. In comparison,
240 48.0% of COVID-19 cases shed ≥1 virion via aerosols in 24 h of talking loudly or singing (*SI*
241 *Appendix*, Fig. S14*C*). Notably, only 61.4% of COVID-19 cases shed ≥1 virion via either
242 droplets or aerosols in 24 h of talking loudly or singing (*SI Appendix*, Fig. S14*D*). While the
243 human infectious dose of SARS-CoV-2 by any exposure route remains unelucidated, it must be
244 at least one viable copy. Thus, at least 38.6% of COVID-19 cases were expected to present
245 negligible risk to spread SARS-CoV-2 through either droplets or aerosols in 24 h. The proportion
246 of inherently infectious cases further decreased as the infectious dose increased: 55.8, 42.5 and
247 25.0% of COVID-19 cases were expected to shed ≥2, ≥10 and ≥100 virions, respectively, in 24 h
248 of talking loudly or singing during the infectious period.

249 While these analyses indicated that a greater proportion of A(H1N1)pdm09 cases were
250 inherently infectious, 18.8% of COVID-19 cases shed virions more rapidly than those infected
251 with A(H1N1)pdm09 (Fig. 4*A*). At the 98th cp for A(H1N1)pdm09, singing expelled 4.38 (2.85-

252 6.78) virions/min and breathing exhaled 2.15×10^{-2} (1.40×10^{-2}-3.34×10^{-2}) virions/min. Highly
253 infectious COVID-19 cases expelled virions at rates that were up to 1-2 orders of magnitude
254 greater than their A(H1N1)pdm09 counterparts (Fig. 5*H* and *SI Appendix*, Fig. S15).
255

256 **Discussion**

257 This study provided systematic analyses of several factors characterizing SARS-CoV-2
258 transmissibility. First, we found that broader heterogeneity in rVL facilitates greater
259 overdispersion for SARS-CoV-2 than A(H1N1)pdm09. Our results suggest that many COVID-
260 19 cases infect no one (12-14) because they inherently present minimal transmission risk via
261 droplets or aerosols, although behavioral and environmental factors can further abate risk.
262 Meanwhile, highly infectious cases can shed tens to thousands of SARS-CoV-2 virions/min,
263 especially between 1-5 DFSO. The model estimates, when corrected to copies rather than
264 virions, align with recent clinical findings for exhalation rates of SARS-CoV-2 (27). In
265 comparison, a greater proportion of A(H1N1)pdm09 cases are infectious but shed virions at low
266 rates, which concurs with more uniform transmission and few superspreading events observed
267 during the 2009 H1N1 pandemic (16, 17). Moreover, our analyses suggest that heterogeneity in
268 rVL may be generally associated with overdispersion for viral respiratory infections. In this case,
269 rVL distribution can serve as an early correlate for transmission patterns, including
270 superspreading, during outbreaks of novel respiratory viruses, providing insight for disease
271 control before large-scale epidemiological studies empirically characterize *k*. When transmission
272 is highly overdispersed, targeted interventions may disproportionately mitigate infection (20),
273 with models estimating that focusing half of control efforts on the most infectious 20% of cases
274 outperforms random control policies threefold (15).

275 Second, we analyzed SARS-CoV-2 kinetics during respiratory infection. While
276 heterogeneity remains broad throughout the infectious period, rVL tends to peak at 1 DFSO and
277 be elevated for 1-5 DFSO, coinciding with the period of highest attack rates observed among
278 close contacts (28). These results indicate that transmission risk tends to be greatest soon after
279 illness rather than in the presymptomatic period, which concurs with large tracing studies (6.4-
280 12.6% of secondary infections from presymptomatic transmission) (29, 30) rather than early
281 temporal models (~44%) (23). Furthermore, our kinetic analysis suggests that, on average,
282 SARS-CoV-2 reaches diagnostic concentrations 1.54-3.17 days after respiratory infection (-3.84
283 to -2.21 DFSO), assuming assay detection limits of 1-3 $\log_{10}$ copies/ml, respectively, for
284 nasopharyngeal swabs immersed in 1 ml of transport media.
285 Third, we assessed the relative infectiousness of COVID-19 subgroups. As a common
286 symptom of COVID-19 (31), coughing sheds considerable numbers of virions via droplets and
287 short-range aerosols. Thus, symptomatic infections tend to be more contagious than
288 asymptomatic ones, providing one reason as to why asymptomatic cases transmit SARS-CoV-2
289 at lower relative rates (32), especially in close contact (33), despite similar rVLs and increased
290 contact patterns. Accordingly, children (48-54% of symptomatic cases present with cough) (34,
291 35) tend to be less contagious than adults (68-80%) (31, 35) based on tendencies of
292 symptomatology rather than rVL. Conversely, coughing sheds few virions via smaller aerosols.
293 Our analyses suggest that asymptomatic and symptomatic infections present comparable risks for
294 airborne spread, as do adult and pediatric cases. While singing and talking loudly, highly
295 infectious cases shed tens to hundreds of SARS-CoV-2 virions/min via long-range and buoyant
296 aerosols.

297 Our study has limitations. The systematic search found a limited number of studies reporting
298 quantitative specimen measurements from the presymptomatic period, meaning these estimates
299 may be sensitive to sampling bias. Although additional studies have reported semiquantitative
300 metrics (cycle thresholds), these data were excluded because they cannot be compared on an
301 absolute scale due to batch effects (36), limiting use in compound analyses. Furthermore, this
302 study considered population-level estimates of the infectious periods, viability proportions and
303 rate profiles for respiratory particles, which omit individual or environmental variation.
304 Distinctions in phonetic tendencies and, especially for young children, respiratory capacity lead
305 to variation in particle emission rates (37). Some patients shed SARS-CoV-2 with diminishing
306 viability soon after symptom onset (21), whereas others produce replication-competent virus for
307 weeks (38). It remains unclear how case characteristics and environmental factors affect the
308 viability dynamics of SARS-CoV-2.
309 Taken together, our findings provide a potential path forward for disease control. They
310 support aerosol spread as a transmission mode for SARS-CoV-2, including for conditional
311 superspreading by highly infectious cases. However, with short durations of stay in well-
312 ventilated areas, the exposure risk for aerosols, including long-range and buoyant ones, remains
313 correlated with proximity to infectious cases (2, 4). Strategies to abate infection should limit
314 crowd numbers and duration of stay while reinforcing distancing, low voice amplitudes and
315 widespread mask usage; well-ventilated settings can be recognized as lower risk venues.
316 Coughing can shed considerable quantities of virions, while rVL tends to peak at 1 DFSO and
317 can be high throughout the infectious period. Thus, immediate, sustained self-isolation upon
318 illness is crucial to curb transmission from symptomatic cases. Collectively, our analyses
319 highlight the role of cases with high rVLs in propelling the COVID-19 pandemic. While

320 diagnosing COVID-19, qRT-PCR can also triage contact tracing, prioritizing these patients: for
321 nasopharyngeal swabs immersed in 1 ml of transport media, ≥7.14 (95% CI: 7.07-7.22) $\log_{10}$
322 copies/ml corresponds to the top 20% of COVID-19 cases. Doing so may identify asymptomatic
323 and presymptomatic infections more efficiently, a key step towards mitigation as the pandemic
324 continues.

Materials and Methods

Systematic review. We undertook a systematic review and prospectively submitted the protocol for registration on PROSPERO (registration number, CRD42020204637). Other than the title of this study, we have followed PRISMA reporting guidelines (39). The systematic review was conducted according to Cochrane methods guidance (40). In-depth details on the search strategy, selection criteria, data collection and contributing studies are included in the *SI Appendix*, Methods.

Calculation of rVLs from specimen measurements. In this study, viral concentrations in respiratory specimens were denoted as specimen measurements, whereas viral concentrations in the respiratory tract were denoted as rVLs. To determine rVLs, each collected quantitative specimen measurement was converted to rVL based on the dilution factor. For example, measurements from swabbed specimens (NPS and OPS) typically report the RNA concentration in viral transport media. Based on the expected uptake volume for swabs (0.128 ± 0.031 ml, mean ± SD) (41) or reported collection volume for expulsed fluid in the study (e.g., 0.5 to 1 ml) along with the reported volume of transport media in the study (e.g., 1 ml), we calculated the dilution factor for each respiratory specimen to estimate the rVL. If the diluent volume was not reported, then the dilution factor was calculated assuming a volume of 1 ml (NPS and OPS), 2 ml (POS and ETA) or 3 ml (NPA) of transport media (42-44). Unless dilution was reported for Spu specimens, we used the specimen measurement as the rVL (21). The non-reporting of diluent volume was noted as an element increasing risk of bias in the hybrid JBI critical appraisal checklist. Specimen measurements (based on instrumentation, calibration, procedures and reagents) are not standardized. While the above procedures (including only quantitative

348 measurements after extraction as an inclusion criterion, considering assay detection limits and
349 correcting for specimen dilution) have considered many of these factors, non-standardization
350 remains an inherent limitation in the variability of specimen measurements.
351
352 **Meta-regression of *k* and heterogeneity in rVL.** To assess the relationship between *k* and
353 heterogeneity in rVL, we performed a univariate meta-regression ($\log k = a(SD) + b$, where a
354 is the slope for association and b is the intercept) between pooled estimates of *k* (based on
355 studies describing community transmission) for COVID-19 (*k* = 0.409) (12-14, 45-48), SARS (*k*
356 = 0.165) (15) and A(H1N1)pdm09 (*k* = 8.155) (16, 17) and the SD of the rVLs in contributing
357 studies. Since SD was the metric, we used a fixed-effects model. For weighting in the meta-
358 regression, we used the proportion of rVL samples from each study relative to the entire
359 systematic dataset ($W_{\mathrm{i}} = n_{\mathrm{i}}/n_{\mathrm{total}}$). All calculations were performed in units of $\log_{10}$ copies/ml.
360 As the meta-regression used pooled estimates of *k* for each infection, it assumed that there was
361 no correlated bias to *k* across contributing studies. The limit of detection for qRT-PCR
362 instruments used in the included studies did not significantly affect the analysis of heterogeneity
363 in rVL, as these limits tended to be below the values found for specimens with low virus
364 concentrations. The meta-regression was conducted using all contributing studies and showed a
365 weak association. Meta-regression was also conducted using studies that had low risk of bias
366 according to the hybrid JBI critical appraisal checklist and showed a strong association. The *P*-
367 value for association was obtained using the meta-regression slope *t*-test for a, the effect
368 estimate.
369

370 **Meta-analysis of rVLs.** Based on the search design and composition of contributing studies, the
371 meta-analysis overall estimates were the expected SARS-CoV-2, SARS-CoV-1 and
372 A(H1N1)pdm09 rVL when encountering a COVID-19, SARS or A(H1N1)pdm09 case,
373 respectively, during their infectious period. Pooled estimates and 95% CIs for the expected rVL
374 of each virus across their infectious period were calculated using a random-effects meta-analysis
375 (DerSimonian and Laird method). The estimates for rVL assumed that each viral copy was
376 extracted and quantified from the tested specimen aliquot. For studies reporting summary
377 statistics in medians and interquartile or total ranges, we derived estimates of the mean and
378 variance and calculated the 95% CIs (49). All calculations were performed in units of $\log_{10}$
379 copies/ml. Between-study heterogeneity in meta-analysis was assessed using Cochran's Q test
380 and the I^2 and τ^2 statistics. If significant between-study heterogeneity in meta-analyses was
381 encountered, sensitivity analysis based on the risk of bias of contributing studies was performed.
382 The meta-analyses were conducted using STATA 14.2 (StataCorp LLC, College Station, Texas,
383 USA).
384
385 **Age and symptomatology subgroup analyses of SARS-CoV-2 rVLs.** The overall estimate for
386 each subgroup was the expected rVL when encountering a case of that subgroup during the
387 infectious period. Studies reporting data exclusively from a subgroup of interest were directly
388 included in the analysis after rVL estimations. For studies in which data for these subgroups
389 constituted only part of its dataset, rVLs from the subgroup were extracted to calculate the mean,
390 variance and 95% CIs. Random-effects meta-analysis was performed as described above. For
391 meta-analyses of pediatric and asymptomatic COVID-19 cases, contributing studies had low risk
392 of bias, and no risk-of-bias sensitivity analyses were performed for these subgroups.

393

394 **Distributions of rVL.** We pooled the entirety of individual sample data in the systematic dataset
395 by disease, COVID-19 subgroups and DFSO. For analyses of SARS-CoV-2 dynamics across
396 disease course, we included estimated rVLs from negative qRT-PCR measurements of
397 respiratory specimens for cases that had previously been quantitatively confirmed to have
398 COVID-19. These rVLs were estimated based on the reported assay detection limit in the
399 respective study. Probability plots and modified Kolmogorov–Smirnov tests used the Blom
400 scoring method and were used to determine the suitability of normal, lognormal, gamma and
401 Weibull distributions to describe the distribution of rVLs for SARS-CoV-2, SARS-CoV-1 and
402 A(H1N1)pdm09. For each virus, the data best conformed to Weibull distributions, which is
403 described by the probability density function

404 $$f(\upsilon) = \frac{\alpha}{\beta}\left(\frac{\upsilon}{\beta}\right)^{\alpha-1} e^{-(\upsilon/\beta)^{\alpha}}, \quad (1)$$

405 where α is the shape factor, β is the scale factor and υ is rVL ($\upsilon \geq 0$ $\log_{10}$ copies/ml). Weibull
406 distributions were fitted on the entirety of collected individual sample data for the respective
407 category. Since individual specimen measurements could not be collected from all studies, there
408 was a small bias on the mean estimate for each fitted distribution. Thus, for the curves shown in
409 Fig. 4, *B* and *C*, the mean of the Weibull distributions summarized in *SI Appendix*, Table S7 was
410 adjusted to be the subgroup meta-analysis estimate for correction; the SD and distribution around
411 that mean remained consistent.

412 For each Weibull distribution, the value of the rVL at the xth percentile was determined
413 using the quantile function,

414 $$\upsilon_x = \beta[-\ln(1-x)]^{1/\alpha}. \quad (2)$$

415 For cp curves, we used eq. (2) to determine rVLs from the 1st cp to the 99th cp (step size, 1%).
416 Curve fitting to eq. (1) and calculation of eq. (2) and its 95% CI was performed using the
417 Distribution Fitter application in Matlab R2019b (MathWorks, Inc., Natick, Massachusetts,
418 USA).
419
420 **Viral kinetics.** To model SARS-CoV-2 kinetics during respiratory infection, we used a
421 mechanistic epithelial cell-limited model for the respiratory tract (50), based on the system of
422 differential equations:

423 $$\frac{dT}{dt} = -\beta TV \quad (3)$$

424 $$\frac{dI}{dt} = \beta TV - \delta I \quad (4)$$

425 $$\frac{dV}{dt} = pI - cV, \quad (5)$$

426 where T is the number of uninfected target cells, I is the number of productively infected cells, V
427 is the rVL, β is the infection rate constant, p is the rate at which airway epithelial cells shed virus
428 to the extracellular fluid, c is the clearance rate of virus and δ is the clearance rate of
429 productively infected cells. Parameter units are summarized in *SI Appendix*, Table S8. Using
430 these parameters, the viral half-life in the respiratory tract ($t_{1/2} = \ln 2/c$) and the half-life of
431 productively infected cells ($t_{1/2} = \ln 2/\delta$) could be estimated. Moreover, the cellular basic
432 reproductive number (the expected number of secondary infected cells from a single
433 productively infected cell placed in a population of susceptible cells) was calculated by

434 $$R_{0,\mathrm{c}} = \frac{p\beta T_0}{c\delta}, \quad (6)$$

435 where T_0 is the initial number of susceptible cells (50). Further details on fitting procedure and
436 analyses of viral kinetics are included in the *SI Appendix*, Methods.
437
438 **Likelihood of respiratory particles containing virions.** To calculate an unbiased estimator for
439 viral partitioning (the expected number of viable copies in an expelled particle at a given size),
440 we multiplied rVLs with the volume equation for spherical particles during atomization and the
441 estimated viability proportion, according to the following equation:

442 $$\lambda = \frac{\pi \rho v_p \gamma \upsilon}{6} d^3, \quad (7)$$

443 where λ is the expectation value, ρ is the material density of the respiratory particle (997 g/m^3),
444 v_p is the volumetric conversion factor (1 ml/g), γ is the viability proportion, υ is the rVL and d is
445 the hydrated diameter of the particle during atomization.
446 The model assumed γ was 0.1% as a population-level estimate. For influenza,
447 approximately 0.1% of copies in particles expelled from the respiratory tract represent viable
448 virus (51), which is equivalent to one in 3 $\log_{10}$ copies/ml for rVL or, after dilution in transport
449 media, roughly one in 4 $\log_{10}$ copies/ml for specimen concentration. Respiratory specimens taken
450 from influenza cases show positive cultures for specimen concentrations down to 4 $\log_{10}$
451 copies/ml (52). Likewise, for COVID-19 cases, recent reports also show culture-positive
452 respiratory specimens with SARS-CoV-2 concentrations down to 4 $\log_{10}$ copies/ml (21),
453 including from pediatric (53) and asymptomatic (22) cases. Our analyses indicated that SARS-
454 CoV-2 rVLs were not different at different sites in the respiratory tract. Moreover, replication-
455 competent SARS-CoV-2 has been found in respiratory specimens taken throughout the
456 respiratory tract (mouth, nasopharynx, oropharynx and lower respiratory tract) (21, 54). Taken
457 together, these considerations suggested that the assumption for viability proportion (0.1%) was

458 suitable to model the likelihood of respiratory particles containing viable SARS-CoV-2. In
459 accordance with the discussion above, the model did not differentiate this population-level
460 viability estimate based on age, symptomatology or sites of atomization. Based on the relative
461 relationship between the residence time of expelled particles before assessment (~5 s) (51), we
462 took the viability proportion (0.1%) to be for equilibrated particles.

463 Likelihood profiles were determined using Poisson statistics, as described by the probability
464 mass function

465 $$\mathrm{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{8}$$

466 where k is the number of virions partitioned within the particle. For λ, 95% CIs were determined
467 using the variance of its rVL estimate. To determine 95% CIs for likelihood profiles from the
468 probability mass function, we used the delta method, which specifies

469 $$\mathrm{Var}\big(g(\boldsymbol{\theta})\big) \approx \sigma^2 \dot{g}(\boldsymbol{\theta})' \mathbf{D} \dot{g}(\boldsymbol{\theta}), \tag{9}$$

470 where $\sigma^2\mathbf{D}$ is the covariance matrix of $\boldsymbol{\theta}$ and $\dot{g}(\boldsymbol{\theta})$ is the gradient of $g(\boldsymbol{\theta})$. For the univariate
471 Poisson distribution, $\sigma^2\mathbf{D} = \lambda$ and

472 $$\dot{g}(\theta) = \frac{\lambda^{k-1} e^{-\lambda}}{k!}(k - \lambda). \tag{10}$$

473

474 **Rate profiles of particles expelled by respiratory activities.** Distributions from the literature
475 were used to determine the rate profiles of particles expelled during respiratory activities. For
476 breathing, talking and coughing, we used data from Johnson et al (55). For singing, we used data
477 from Morawaska et al (56) for smaller aerosols (d_a < 20 μm) and used the profiles from talking
478 for larger aerosols and droplets based on the oral cavity mechanism from Johnson et al (55).
479 Further details on the calculation of rate profiles are included in the *SI Appendix*, Methods. We

480 compared these rate profiles with those collected from talking loudly and talking quietly from
481 Asadi et al (57). After dehydration, the particle diameter becomes approximately 0.5 times the
482 initial size from when atomized in the respiratory tract (55), which was used as the factor
483 between the diameter of hydrated particles during atomization and dehydrated particles.
484 Equilibrium aerodynamic diameter was calculated by $d_a = d_p(\rho/\rho_0)^{1/2}$, where d_p is the
485 dehydrated diameter, ρ is the material density of the respiratory particle (taken to be 1 g/cm^3
486 based on the composition of dehydrated respiratory particles) and ρ_0 is the reference material
487 density (1 g/cm^3). Curves based on discrete particle measurements were connected using the
488 nonparametric Akima spline function.
489
490 **Shedding virions via respiratory droplets and aerosols.** To model the respiratory shedding
491 rate across particle size, rVL estimates and the hydrated diameters of particles expelled by a
492 respiratory activity were input into eq. (7), and the output was then multiplied by the rate profile
493 of the activity (talking, singing, breathing or coughing). To assess the relative contribution of
494 aerosols and droplets to mediating respiratory viral shedding for a given respiratory activity, we
495 calculated the proportion of the cumulative hydrated volumetric rate contributed by buoyant
496 aerosols ($d_a \leq 10$ μm), long-range aerosols (10 μm < $d_a \leq 50$ μm), short-range aerosols (50 μm <
497 $d_a \leq 100$ μm) and droplets ($d_a > 10$ μm) for that respiratory activity. Since the Poisson mean was
498 proportional to cumulative volumetric rate, this estimate of the relative contribution of aerosols
499 and droplets to respiratory viral shedding was consistent among viruses and cps in the model.
500 To determine the total respiratory shedding rate for a given respiratory activity across cp, we
501 determined the cumulative hydrated volumetric rate (by summing the hydrated volumetric rates
502 across particle sizes for that respiratory activity) of particle atomization and input it into eq. (7).

503 Using rVLs and their variances as determined by the Weibull quantile functions, we then
504 calculated the Poisson means and their 95% CIs at the different cps.
505 To assess the influence of heterogeneity in rVL on individual infectiousness, we first
506 considered transmission of A(H1N1)pdm09 via aerosols (19). The 50% human infectious dose
507 (HID_{50}) of aerosolized A(H1N1)pdm09 was taken to be 1-3 virions (26). To determine the
508 expected time required for a A(H1N1)pdm09 case to shed 1 virion via aerosols, we took the
509 reciprocal of the Poisson means and their 95% CIs at the different cps of the estimated shedding
510 rates. The expected time required for a COVID-19 case to shed 1 virion via aerosols or 1 virion
511 via droplets or aerosols was determined in a same manner.
512
513 **Data availability.** Data will be made available upon request. All raw data, code and model
514 outputs from this study will be made publicly available in online repositories after peer review.
515 Search strategies for the systematic review are shown in Supplementary Tables 1-5. The
516 systematic review protocol was prospectively registered on PROSPERO (registration number,
517 CRD42020204637).
518
519 **Acknowledgments**
520 We thank T. Alba (Toronto) for discussion on statistical methods. We thank J. Jimenez
521 (Colorado) for discussion on the characteristics of aerosols and droplets. We thank E. Lavezzo
522 and A. Chrisanti (Padova) and A. Wyllie, A. Ko and N. Grubaugh (Yale) for responses to data
523 inquiries. This work was supported by the Natural Sciences and Engineering Research Council
524 of Canada (NSERC) and the Toronto COVID-19 Action Fund. P.Z.C. was supported by the
525 NSERC Vanier Scholarship (608544). D.N.F. was supported by the Canadian Institutes of Health

Research (Canadian COVID-19 Rapid Research Fund, OV4-170360). F.X.G. was supported by the NSERC Senior Industrial Research Chair.

Author Contributions: P.Z.C. designed research, analyzed data and wrote the paper. P.Z.C. and N.B. conducted screening, appraised studies and drafted the review protocol. Z.P. developed and conducted the systematic review search. M.K. and D.N.F. assessed methods, interpreted results and contributed to the discussion. All authors reviewed and revised the manuscript. F.X.G. supervised the research.

Competing Interest Statement: The authors declare no competing interest.

536 **References**

537 1. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in
538 real time. *Lancet Infect. Dis.* **20**, 533-534 (2020).
539 2. K. A. Prather *et al.*, Airborne transmission of SARS-CoV-2. *Science* **370**, 303-304
540 (2020).
541 3. C. J. Roy, D. K. Milton, Airborne transmission of communicable infection — the elusive
542 pathway. *N. Engl. J. Med.* **350**, 1710-1712 (2004).
543 4. L. Liu, Y. Li, P. V. Nielsen, J. Wei, R. L. Jensen, Short-range airborne transmission of
544 expiratory droplets between two people. *Indoor Air* **27**, 452-462 (2017).
545 5. J. J. Wei, Y. G. Li, Enhanced spread of expiratory droplets by turbulence in a cough jet.
546 *Build. Environ.* **93**, 86-96 (2015).
547 6. D. K. Chu *et al.*, Physical distancing, face masks, and eye protection to prevent person-
548 to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-
549 analysis. *Lancet* **395**, 1973-1987 (2020).
550 7. S. Y. Park *et al.*, Coronavirus disease outbreak in call center, South Korea. *Emerg. Infect.*
551 *Dis.* **26**, 1666-1670 (2020).
552 8. J. Y. Lu *et al.*, COVID-19 outbreak associated with air conditioning in restaurant,
553 Guangzhou, China, 2020. *Emerg. Infect. Dis.* **26**, 1628-1631 (2020).
554 9. L. Hamner *et al.*, High SARS-CoV-2 attack rate following exposure at a choir practice -
555 Skagit County, Washington, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 606-610
556 (2020).
557 10. Q. Li *et al.*, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected
558 pneumonia. *N. Engl. J. Med.* **382**, 1199-1207 (2020).

559 11. X. Hao *et al.*, Reconstruction of the full transmission dynamics of COVID-19 in Wuhan.
560 *Nature* **584**, 420-424 (2020).
561 12. R. Laxminarayan *et al.*, Epidemiology and transmission dynamics of COVID-19 in two
562 Indian states. *Science* **370**, 691-697 (2020).
563 13. A. Endo, S. Abbott, A. J. Kucharski, S. Funk, Estimating the overdispersion in COVID-
564 19 transmission using outbreak sizes outside China. *Wellcome Open Research* **5** (2020).
565 14. Q. Bi *et al.*, Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their
566 close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious*
567 *Diseases* **20**, 911-919 (2020).
568 15. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect
569 of individual variation on disease emergence. *Nature* **438**, 355-359 (2005).
570 16. M. G. Roberts, H. Nishiura, Early estimation of the reproduction number in the presence
571 of imported cases: pandemic influenza H1N1-2009 in New Zealand. *PLoS One* **6**, e17835
572 (2011).
573 17. J. Brugger, C. L. Althaus, Transmission of and susceptibility to seasonal influenza in
574 Switzerland from 2003 to 2015. *Epidemics* **30**, 100373 (2020).
575 18. I. T. Yu *et al.*, Evidence of airborne transmission of the severe acute respiratory
576 syndrome virus. *N. Engl. J. Med.* **350**, 1731-1739 (2004).
577 19. B. J. Cowling *et al.*, Aerosol transmission is an important mode of influenza A virus
578 spread. *Nature Communications* **4**, 1935 (2013).
579 20. E. C. Lee, N. I. Wada, M. K. Grabowski, E. S. Gurley, J. Lessler, The engines of SARS-
580 CoV-2 spread. *Science* **370**, 406-407 (2020).

581 21. R. Wolfel *et al.*, Virological assessment of hospitalized patients with COVID-2019.
582 *Nature* **581**, 465-469 (2020).
583 22. M. M. Arons *et al.*, Presymptomatic SARS-CoV-2 infections and transmission in a
584 skilled nursing facility. *N. Engl. J. Med.* **382**, 2081-2090 (2020).
585 23. X. He *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19.
586 *Nat. Med.* **26**, 672-675 (2020).
587 24. V. E. Pitzer, G. M. Leung, M. Lipsitch, Estimating variability in the transmission of
588 severe acute respiratory syndrome to household contacts in Hong Kong, China. *Am. J.*
589 *Epidemiol.* **166**, 355-363 (2007).
590 25. D. K. Ip *et al.*, Viral shedding and transmission potential of asymptomatic and
591 paucisymptomatic influenza virus infections in the community. *Clin. Infect. Dis.* **64**, 736-
592 742 (2017).
593 26. P. Fabian *et al.*, Influenza virus in human exhaled breath: an observational study. *PLoS*
594 *One* **3**, e2691 (2008).
595 27. J. Ma *et al.*, COVID-19 patients in earlier stages exhaled millions of SARS-CoV-2 per
596 hour. *Clin. Infect. Dis.* 10.1093/cid/ciaa1283 (2020).
597 28. H. Y. Cheng *et al.*, Contact tracing assessment of COVID-19 transmission dynamics in
598 Taiwan and risk at different exposure periods before and after symptom onset. *JAMA*
599 *Intern Med* **180**, 1156-1163 (2020).
600 29. W. E. Wei *et al.*, Presymptomatic transmission of SARS-CoV-2 — Singapore, January
601 23-March 16, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 411-415 (2020).
602 30. Z. Du *et al.*, Serial interval of COVID-19 among publicly reported confirmed cases.
603 *Emerg. Infect. Dis.* **26**, 1341-1343 (2020).

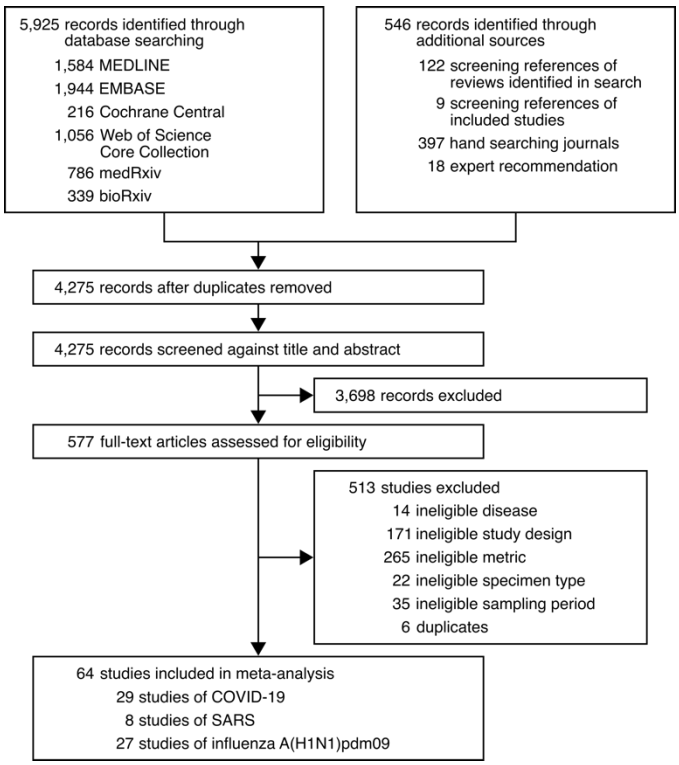
604 31. W. J. Guan *et al.*, Clinical characteristics of coronavirus disease 2019 in China. *N. Engl.*
605 *J. Med.* **382**, 1708-1720 (2020).
606 32. R. Li *et al.*, Substantial undocumented infection facilitates the rapid dissemination of
607 novel coronavirus (SARS-CoV-2). *Science* **368**, 489-493 (2020).
608 33. L. Luo *et al.*, Contact settings and risk for transmission in 3410 close contacts of patients
609 with COVID-19 in Guangzhou, China: a prospective cohort study. *Ann. Intern. Med.*
610 10.7326/M20-2671 (2020).
611 34. X. Lu *et al.*, SARS-CoV-2 infection in children. *N. Engl. J. Med.* **382**, 1663-1665 (2020).
612 35. C. C.-R. Team, Coronavirus disease 2019 in children - United States, February 12-April
613 2, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 422-426 (2020).
614 36. M. S. Han, J. H. Byun, Y. Cho, J. H. Rim, RT-PCR for SARS-CoV-2: quantitative versus
615 qualitative. *Lancet Infect. Dis.* 10.1016/S1473-3099(20)30424-2 (2020).
616 37. H.-L. Yen *et al.*, Effect of voicing and articulation manner on aerosol particle emission
617 during human speech. *PLoS One* **15** (2020).
618 38. J. J. A. van Kampen *et al.*, Shedding of infectious virus in hospitalized patients with
619 coronavirus disease-2019 (COVID-19): duration and key determinants.
620 medRxiv:10.1101/2020.06.08.20125310 (2020).
621 39. D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, P. Group, Preferred reporting items for
622 systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* **6**, e1000097
623 (2009).
624 40. J. P. T. Higgins *et al.* Cochrane handbook for systematic reviews of interventions. in
625 *Cochrane book series* (John Wiley & Sons, 2019).

626 41. P. Warnke, L. Warning, A. Podbielski, Some are more equal - a comparative study on
627 swab uptake and release of bacterial suspensions. *PLoS One* **9**, e102215 (2014).
628 42. E. Lavezzo *et al.*, Suppression of a SARS-CoV-2 outbreak in the Italian municipality of
629 Vo'. *Nature* **584**, 425-429 (2020).
630 43. K. K. To *et al.*, Temporal profiles of viral load in posterior oropharyngeal saliva samples
631 and serum antibody responses during infection by SARS-CoV-2: an observational cohort
632 study. *Lancet Infect. Dis.* **20**, 565-574 (2020).
633 44. L. L. Poon *et al.*, Detection of SARS coronavirus in patients with severe acute respiratory
634 syndrome by conventional and real-time quantitative reverse transcription-PCR assays.
635 *Clin. Chem.* **50**, 67-72 (2004).
636 45. D. C. Adam *et al.*, Clustering and superspreading potential of SARS-CoV-2 infections in
637 Hong Kong. *Nat. Med.* 10.1038/s41591-020-1092-0 (2020).
638 46. J. Riou, C. L. Althaus, Pattern of early human-to-human transmission of Wuhan 2019
639 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.* **25**
640 (2020).
641 47. Y. Zhang, Y. Li, L. Wang, M. Li, X. Zhou, Evaluating transmission heterogeneity and
642 super-spreading event of COVID-19 in a metropolis of China. *Int. J. Environ. Res. Public*
643 *Health* **17**, 3705 (2020).
644 48. A. Tariq *et al.*, Real-time monitoring the transmission potential of COVID-19 in
645 Singapore, March 2020. *BMC Med.* **18**, 166 (2020).
646 49. X. Wan, W. Wang, J. Liu, T. Tong, Estimating the sample mean and standard deviation
647 from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.*
648 **14**, 135 (2014).

649 50. P. Baccam, C. Beauchemin, C. A. Macken, F. G. Hayden, A. S. Perelson, Kinetics of
650 influenza A virus infection in humans. *J. Virol.* **80**, 7590-7599 (2006).
651 51. J. Yan *et al.*, Infectious virus in exhaled breath of symptomatic seasonal influenza cases
652 from a college community. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 1081-1086 (2018).
653 52. L. L. Lau *et al.*, Viral shedding and clinical illness in naturally acquired influenza virus
654 infections. *J. Infect. Dis.* **201**, 1509-1516 (2010).
655 53. A. G. L'Huillier, G. Torriani, F. Pigny, L. Kaiser, I. Eckerle, Culture-competent SARS-
656 CoV-2 in nasopharynx of symptomatic neonates, children, and adolescents. *Emerg.*
657 *Infect. Dis.* **26**, 2494-2497 (2020).
658 54. H. W. Jeong *et al.*, Viable SARS-CoV-2 in various specimens from COVID-19 patients.
659 *Clin. Microbiol. Infect.* 10.1016/j.cmi.2020.07.020 (2020).
660 55. G. R. Johnson *et al.*, Modality of human expired aerosol size distributions. *J. Aerosol Sci*
661 **42**, 839-851 (2011).
662 56. L. Morawska *et al.*, Size distribution and sites of origin of droplets expelled from the
663 human respiratory tract during expiratory activities. *J. Aerosol Sci* **40**, 256-269 (2009).
664 57. S. Asadi *et al.*, Aerosol emission and superemission during human speech increase with
665 voice loudness. *Sci. Rep.* **9**, 2348 (2019).

666 **Figures**

667



668

669

670 **Fig. 1.** Development of the systematic dataset.

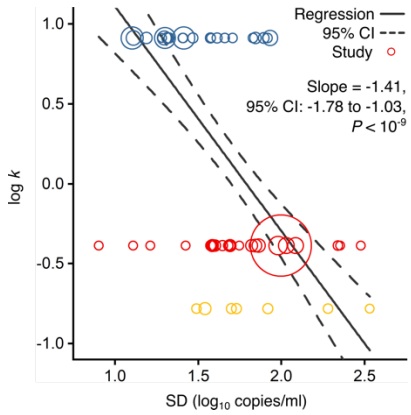
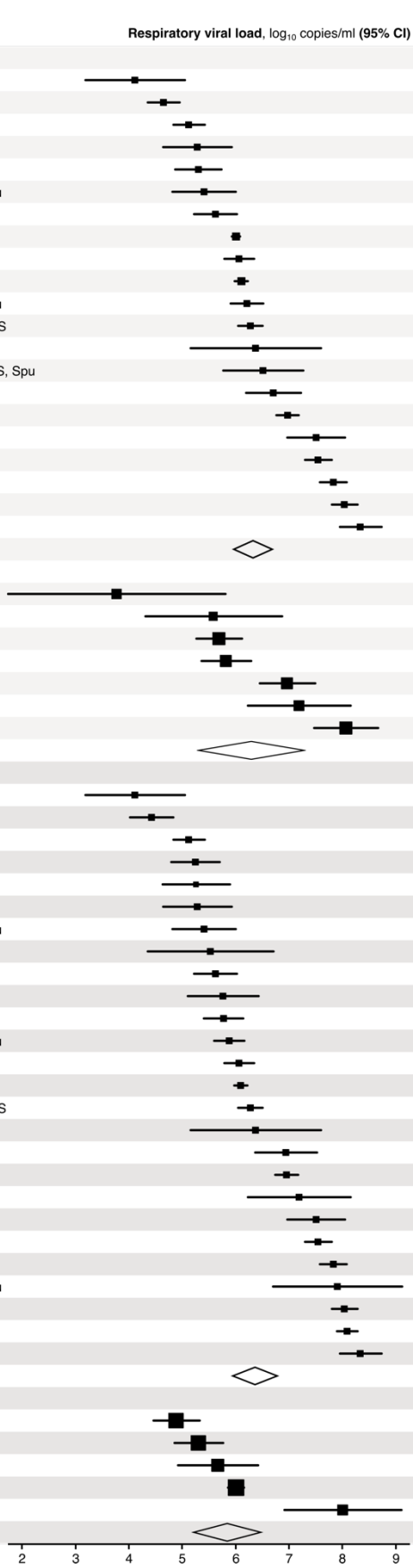


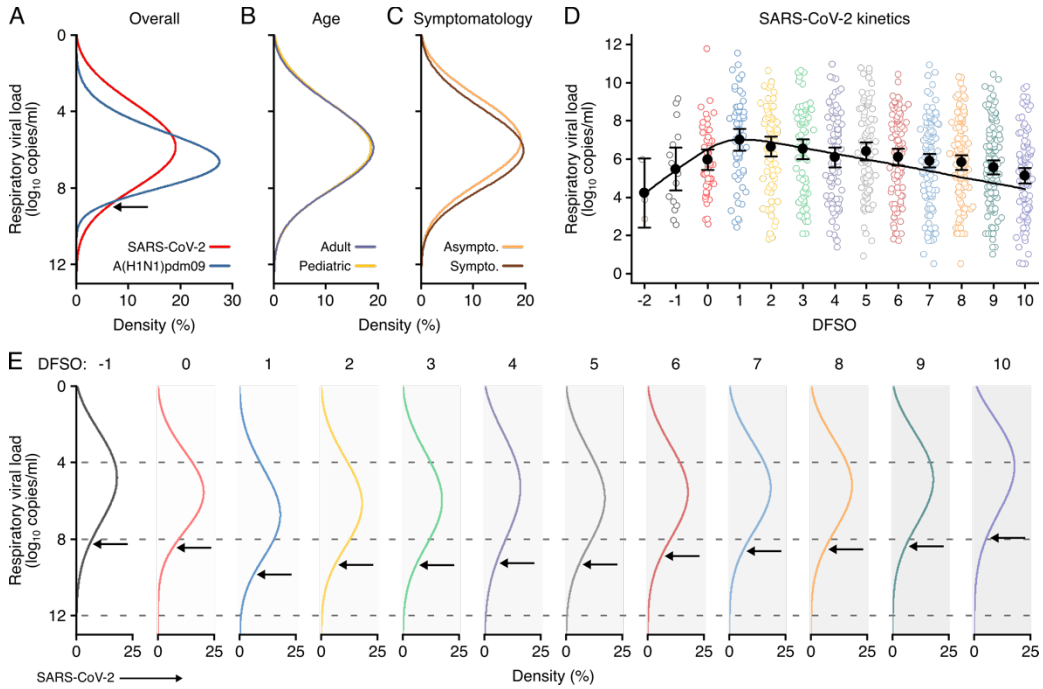
Fig. 2. Association of overdispersion in SARS-CoV-2, SARS-CoV-1 and A(H1N1)pdm09 transmissibility with heterogeneity in rVL. Meta-regression of dispersion parameter (*k*) with the standard deviation (SD) of respiratory viral loads (rVLs) from contributing studies with low risk of bias (Pearson's $r = -0.73$). Pooled estimates of k were determined from the literature for each infection. Blue, red and yellow circles denote A(H1N1)pdm09 ($N = 22$), COVID-19 ($N = 24$) and SARS ($N = 7$) studies, respectively. Circle sizes denote weighting in the meta-regression. The *P*-value was obtained using the meta-regression slope *t*-test.

Subgroup, study ref.	Cases No.	Cases Type	Specimens No.	Specimens DFSO	Specimens Type	Respiratory viral load, log_{10} copies/ml (95% CI)	*P* value for difference
Adult COVID-19							
62	4	H, A, S	9	4 to 10	NPS, OPS	4.12 (3.18-5.05)	
42	100	H, C, A, S, Ps, As	106	-2 to 10	NPS, OPS	4.65 (4.35-4.95)	
63	6	H, A, S	6	3 to 8	OPS	5.13 (4.83-5.42)	
43	23	H, A, S	51	0 to 10	ETA, POS	5.29 (4.65-5.93)	
64	14	H, A, S, As	55	0 to 10	NPS, OPS	5.30 (4.87-5.74)	
65	-	H, N, A, S	31	2 to 10	NPS, OPS, Spu	5.41 (4.81-6.01)	
66	-	H, A, S	19	1 to 10	POS, Spu	5.62 (5.22-6.03)	
67	2,182	C, A, As	2,182	0	NPS	6.01 (5.93-6.09)	
68	171	N, A, S	171	0 to 10	NPS	6.07 (5.79-6.34)	
69	103	H, N, A, S	103	0 to 8	NPS, OPS	6.10 (5.98-6.23)	
21	9	H, A, S, As	133	2 to 10	NPS, OPS, Spu	6.21 (5.90-6.52)	
70	41	H, A, S	310	2 to 10	NPS, OPS, POS	6.28 (6.05-6.50)	
71	5	H, A, S	5	7 to 9	NPS	6.38 (5.16-7.60)	
72	-	H, A, S, As	38	0 to 9	NPA, NPS, OPS, Spu	6.52 (5.77-7.27)	
73	39	H, A, S, As	41	0 to 10	NPS	6.71 (6.20-7.23)	
74	347	H, A, S	347	0 to 5	NPS	6.97 (6.77-7.18)	
75	24	H, A, S	33	0 to 10	NPS	7.51 (6.97-8.05)	
76	205	H, N, A, S	205	0 to 10	NPS	7.55 (7.30-7.81)	
38	-	H, A, S	154	0 to 10	NPS, Spu	7.83 (7.58-8.08)	
77	148	N, A, S	296	1 to 8	NPS	8.04 (7.81-8.28)	
78	5	H, A, S	63	0 to 10	NPS, OPS	8.34 (7.95-8.73)	
Overall ($Q_{20} > 100$, $P < 0.001$, $I^2 = 97.7\%$)						6.33 (5.95-6.70)	-
Pediatric COVID-19							
42	2	H, C, P, S	3	8 to 9	NPS, OPS	3.77 (1.74-5.81)	
79	14	H, P, S	14	0 to 10	NPS	5.59 (4.31-6.87)	
80	7	H, P, S, Ps, As	14	-1 to 10	NPS	5.69 (5.26-6.12)	
81	54	C, P, S, Ps, As	54	-3 to 10	NPS	5.82 (5.36-6.29)	
74	58	H, P, S	58	0 to 5	NPS	6.97 (6.45-7.49)	
53	23	H, P, S	23	0 to 4	NPS	7.20 (6.23-8.16)	
82	12	H, P, S, As	27	0 to 10	NPS	8.07 (7.47-8.67)	
Overall ($Q_6 = 20.5$, $P = 0.002$, $I^2 = 79.7\%$)						6.29 (5.27-7.32)	0.476
Symptomatic/presymptomatic COVID-19							
62	4	H, A, S	9	4 to 10	NPS, OPS	4.12 (3.18-5.05)	
42	53	H, C, A, P, S, Ps	60	-2 to 10	NPS, OPS	4.43 (4.03-4.84)	
63	6	H, A, S	6	3 to 8	OPS	5.13 (4.83-5.42)	
64	13	H, A, S	51	0 to 10	NPS, OPS	5.25 (4.79-5.71)	
83	77	H, S, Ps	119	-1 to 8	OPS, Spu	5.26 (4.63-5.90)	
43	23	H, A, S	51	0 to 10	ETA, POS	5.29 (4.65-5.93)	
65	-	H, N, A, S	31	2 to 10	NPS, OPS, Spu	5.41 (4.81-6.01)	
79	17	H, A, P, S	17	0 to 10	NPS	5.53 (4.36-6.71)	
66	-	H, A, S	19	1 to 10	POS, Spu	5.62 (5.22-6.03)	
80	6	H, P, S, Ps	13	-1 to 10	NPS	5.77 (5.10-6.53)	
81	81	C, A, P, S, Ps	81	-3 to 10	NPS	5.78 (5.41-6.15)	
21	8	H, A, S	129	2 to 10	NPS, OPS, Spu	5.88 (5.60-6.17)	
68	171	N, A, S	171	0 to 10	NPS	6.07 (5.79-6.34)	
69	103	H, N, A, S	103	0 to 8	NPS, OPS	6.10 (5.98-6.23)	
70	41	H, A, S	310	2 to 10	NPS, OPS, POS	6.28 (6.05-6.50)	
71	5	H, A, S	5	7 to 9	NPS	6.38 (5.16-7.60)	
73	31	H, A, P, S	33	2 to 10	NPS	6.95 (6.37-7.53)	
74	405	H, A, P, S	405	0 to 5	NPS	6.95 (6.78-7.17)	
53	23	H, P, S	23	0 to 4	NPS	7.20 (6.23-8.16)	
75	24	H, A, S	33	0 to 10	NPS	7.51 (6.97-8.05)	
76	205	H, N, A, S	205	0 to 10	NPS	7.55 (7.30-7.81)	
38	-	H, A, S	154	0 to 10	NPS, Spu	7.83 (7.58-8.08)	
84	2	H, A, P, S	8	4 to 10	NPS, OPS, Spu	7.91 (6.69-9.12)	
77	148	N, A, S	296	1 to 8	NPS	8.04 (7.81-8.28)	
82	9	H, P, S	20	0 to 10	NPS	8.10 (7.37-8.33)	
78	5	H, A, S	63	0 to 10	NPS, OPS	8.34 (7.95-8.73)	
Overall ($Q_{25} > 100$, $P < 0.001$, $I^2 = 91.1\%$)						6.37 (5.93-6.80)	-
Asymptomatic COVID-19							
42	49	H, C, A, As	49	0	NPS, OPS	4.89 (4.46-5.33)	
81	52	C, A, P, As	52	0	NPS	5.31 (4.85-5.77)	
73	9	H, A, As	9	0	NPS	5.67 (4.92-6.43)	
67	2,200	C, A, P, As	2,200	0	NPS	6.01 (5.93-6.09)	
82	3	H, P, As	7	0 to 8	NPS	8.01 (6.91-9.11)	
Overall ($Q_4 = 47.1$, $P < 0.001$, $I^2 = 91.5\%$)						5.84 (5.19-6.50)	0.090



681

682 **Fig. 3.** Subgroup analyses of SARS-CoV-2 rVL during the infectious period. Random-effects
683 meta-analyses comparing the expected respiratory viral loads (rVLs) of adult (≥18 years old)
684 COVID-19 cases with pediatric (<18 years old) ones (top) and symptomatic/presymptomatic
685 infections with asymptomatic ones (bottom) during the infectious period. Quantitative rVLs refer
686 to virus concentrations in the respiratory tract. Case types: hospitalized (H), not admitted (N),
687 community (C), adult (A), pediatric (P), symptomatic (S), presymptomatic (Ps) and
688 asymptomatic (As). Specimen types: endotracheal aspirate (ETA), nasopharyngeal aspirate
689 (NPA), nasopharyngeal swab (NPS), oropharyngeal swab (OPS), posterior oropharyngeal saliva
690 (POS) and sputum (Spu). Dashes denote case numbers that were not obtained. Box sizes denote
691 weighting in the overall estimates. Between-study heterogeneity was assessed using the *P*-value
692 from Cochran's *Q* test and the I^2 statistic. References after 57 are listed in *SI Appendix*,
693 References. One-sided Welch's *t*-tests compared expected rVLs between the COVID-19
694 subgroups (non-significance, $P > 0.05$).



695

696

697 **Fig. 4. Heterogeneity and kinetics of SARS-CoV-2 rVL.** (*A*) Estimated distribution of
698 respiratory viral load (rVL) for SARS-CoV-2 (N = 3,834 samples from N = 26 studies) and
699 A(H1N1)pdm09 (N = 512 samples from N = 10 studies) throughout the infectious periods. (*B*
700 and *C*) Estimated distribution of SARS-CoV-2 rVL for adult (N = 3,575 samples from N = 20
701 studies) and pediatric (N = 198 samples from N = 9 studies) (*B*) and
702 symptomatic/presymptomatic (N = 1,574 samples from N = 22 studies) and asymptomatic (N =
703 2,221 samples from N = 7 studies) (*C*) COVID-19 cases. (*D*) SARS-CoV-2 rVLs fitted to a
704 mechanistic model of viral kinetics (black curve, r^2 = 0.84 for mean estimates). Filled circles and
705 bars depict mean estimates and 95% CIs. Open circles show the entirety of individual sample
706 data over days from symptom onset (DFSO) (left to right, N = 3, 15, 50, 63, 71, 75, 85, 93, 105,
707 136, 123, 128 and 115 samples from N = 21 studies). (*E*) Estimated distributions of SARS-CoV-
708 2 rVL across DFSO. Weibull distributions were fitted on the entirety of individual sample data

709 for the virus, subgroup or DFSO in the systematic dataset. Arrows denote 90th case percentiles

710 for SARS-CoV-2 rVL distributions.

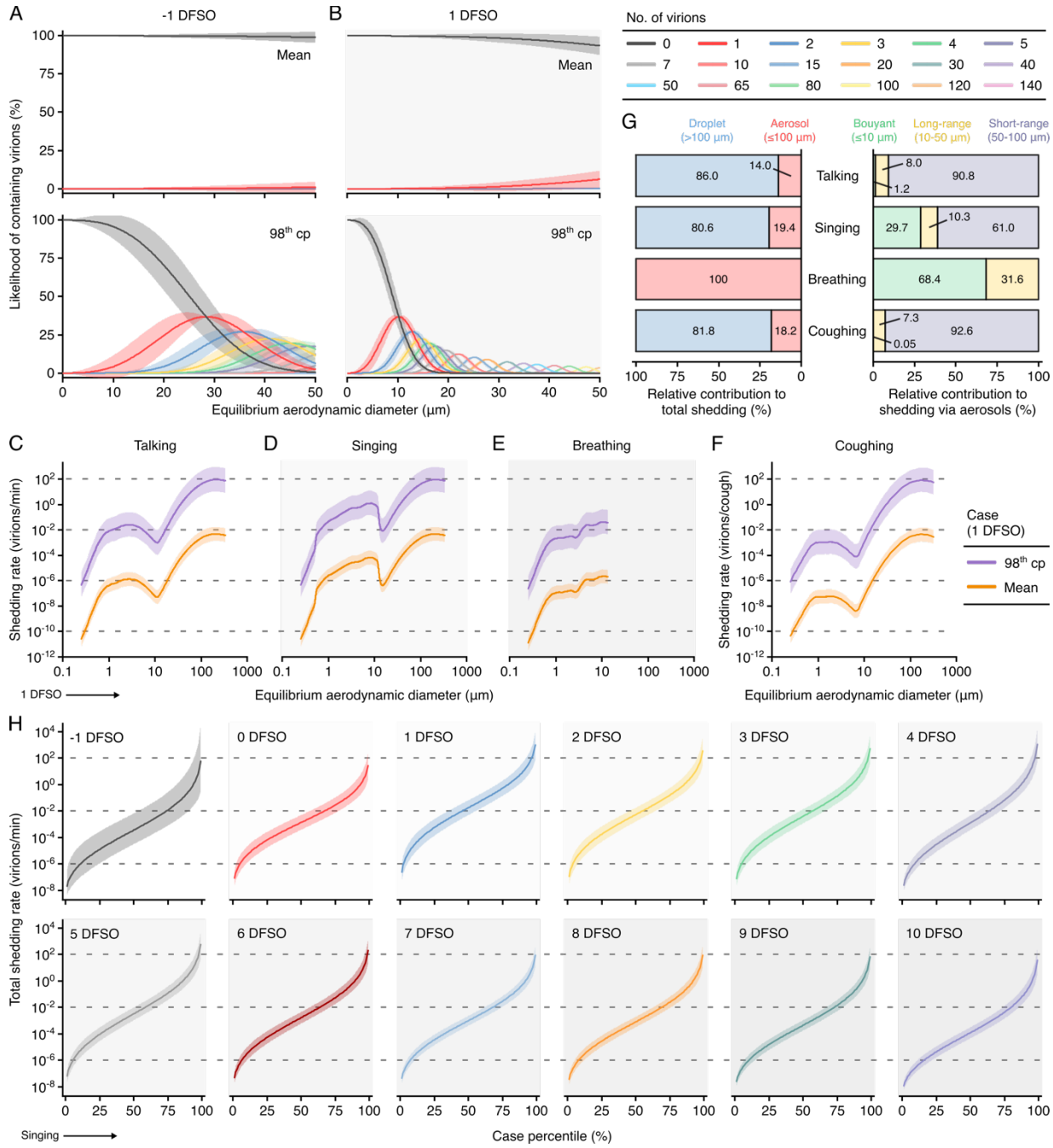


Fig. 5. Heterogeneity in shedding SARS-CoV-2 via droplets and aerosols. (*A* and *B*) Estimated likelihood of respiratory particles containing viable SARS-CoV-2 when expelled by the mean (top) or 98th case percentile (cp) (bottom) COVID-19 cases at -1 (*A*) or 1 (*B*) days from symptom onset (DFSO). For higher no. of virions, some likelihood curves were omitted to aid

717 visualization. When the likelihood for 0 virions approaches 0%, particles are expected to contain
718 at least one viable copy. (*C* to *F*) Rate that the mean and 98th-cp COVID-19 cases at 1 DFSO
719 shed viable SARS-CoV-2 by talking, singing, breathing or coughing over particle size. (*G*)
720 Relative contributions of droplets and aerosols to shedding virions for each respiratory activity
721 (left). Relative contribution of buoyant, long-range and short-range aerosols to shedding virions
722 via aerosols for each respiratory activity (right). (*H*) Case heterogeneity in the total shedding rate
723 (over all particle sizes) of virions via singing across the infectious period. Earlier
724 presymptomatic days were excluded based on limited data. Data range between the 1st and 99th
725 cps. Lines and bands represent estimates and 95% CIs, respectively, for estimated likelihoods or
726 Poisson means.