

21 **SUMMARY**

22 SARS-CoV-2 superspreading occurs when transmission is highly efficient and/or an
23 individual infects many others, contributing to rapid spread. To better quantify heterogeneity in
24 SARS-CoV-2 transmission, particularly superspreading, we performed a systematic review of
25 transmission events with data on secondary attack rates or contact tracing of individual index
26 cases published before September 2021, prior to emergence of variants of concern and
27 widespread vaccination. We reviewed 592 distinct events and 9,883 index cases from 491
28 papers. Meta-analysis of secondary attack rates identified substantial heterogeneity across 12
29 chosen event types/settings, with the highest transmission (25–35%) in co-living situations
30 including households, nursing homes, and other congregate housing. Among index cases, 67%
31 produced zero secondary cases and only 3% (287) infected >5 secondary cases
32 (“superspreaders”). Index case demographic data was limited, with only 55% of individuals
33 reporting age, sex, symptoms, real-time PCR cycle threshold values, or total contacts. With the
34 data available, we identified a higher percentage of superspreaders among symptomatic
35 individuals, individuals aged 49–64 years, and individuals with over 100 total contacts.
36 Addressing gaps in reporting on transmission events and contact tracing in the literature is
37 needed to properly explain heterogeneity in transmission and facilitate control efforts for SARS-
38 CoV-2 and other infections.

39

40 **KEYWORDS**

41 coronavirus; COVID-19; transmission; heterogeneity; infectious disease epidemiology

42

43 INTRODUCTION

44 Following the emergence of SARS-CoV-2 in late 2019, the virus spread worldwide,
45 resulting in the coronavirus disease (COVID-19) pandemic [1]. Understanding drivers of SARS-
46 CoV-2 transmission was crucial for formulating control measures, especially prior to the
47 development of vaccines. Early in the pandemic, heterogeneity in transmission, particularly
48 superspreading, was investigated because of its ability to cause large outbreaks [2–4].
49 Superspreading involves two distinct but non-mutually exclusive phenomena: a setting where
50 many people become infected due to an environment conducive to transmission (e.g., crowded
51 indoor settings), and individuals who are outliers in the number of secondary cases they infect,
52 due to biological heterogeneity in infectiousness and/or engagement in high-risk behaviors [5,6].
53 Superspreading has been observed in several other viral infections, including SARS-CoV,
54 MERS-CoV, Nipah, Ebola, and measles [7–12]. With SARS-CoV-2, both forms of
55 superspreading garnered considerable attention in the literature. For example, over 140
56 individuals were infected during a Christmas event in Belgium in December 2020, causing over
57 26 deaths [13]. Likewise, one individual infected dozens of people during a choir practice in
58 Washington, USA, in March 2020 [14].

59 Because superspreading events contributed substantially to local and global SARS-CoV-
60 2 transmission [15], public health interventions were enacted to reduce their risk of occurrence.
61 These interventions included school closures, limitations on indoor gatherings, and restrictions
62 on visiting hospitalized patients or long-term care facilities. Many of these policies were based
63 on limited data from early in the pandemic. Moreover, published systematic reviews and
64 modeling of SARS-CoV-2 superspreading from this period were limited in scope and did little to
65 disaggregate this phenomenon into the distinct contributions of environment and individual

66 characteristics. For example, studies of setting-specific transmission rates have focused on
67 household and healthcare transmission or geographic and temporal trends [2,16–19], but did not
68 address transmission heterogeneity across other social settings. Previous meta-analyses of
69 individual-level superspreading included only a small number of papers (<26) that calculated
70 overdispersion in transmission, missing the majority of published transmission trees and
71 capturing data primarily from Asia [7,8]. Early investigations of individual-level characteristics
72 related to superspreading were also limited by incomplete contact tracing [20,21] and a focus on
73 clinical over demographic characteristics [20]. A more complete summary of superspreading is
74 needed to understand the scale of transmission heterogeneity across settings and identify causes
75 of individual heterogeneity.

76 The objective of this review was to summarize global heterogeneity in SARS-CoV-2
77 transmission events prior to widespread vaccination and the role of environmental and individual
78 factors in superspreading. Specifically, this review aimed to identify: 1) the amount of variation
79 in attack rates across studies and events, 2) which settings had the highest attack rates, 3) the
80 individual offspring distribution for SARS-CoV-2, and 4) the characteristics of superspreading
81 individuals.

82

83 **METHODS**

84 **Literature search and data extraction**

85 We conducted this systematic review and meta-analysis according to the Preferred
86 Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [22]; see
87 Appendix 1 for the PRISMA checklist. We included all studies of SARS-CoV-2 that contained
88 data on: 1) transmission chains; 2) numbers of index cases, contacts, and infected contacts; 3)

89 numbers of index cases and infected contacts; or 4) secondary attack rates, i.e., number of
90 infected contacts divided by number of contacts. We excluded studies that were not about
91 humans. A clinical informationist searched PubMed, the WHO COVID database, the I Love
92 Evidence COVID database, and Embase on 9 September 2021. No restrictions on language or
93 start date were applied. Results were imported into EndNote X9 (Clarivate, London, UK) where
94 duplicates with exact matches in the author, year, and title fields were removed. Team members
95 screened titles and abstracts and performed full text review in Covidence (Veritas Health
96 Innovation, Melbourne, Australia).

97 We extracted data using a pre-designed, study-specific spreadsheet, collecting
98 information on paper metadata and target variables for two outcomes: transmission events and
99 individual index cases (Table 1). Events were defined as discrete transmission events where
100 secondary attack rates for defined groups of people could be calculated as the number of infected
101 cases divided by the total number of exposed individuals. This definition of secondary attack
102 rates includes both clinical and subclinical infections in some studies. Due to the limited details
103 published in the literature, we did not attempt to distinguish events associated with individual
104 transmission chains from a single source (potentially with confirmatory sequencing data) from
105 events that aggregated multiple transmission chains together. In lieu of this distinction, we
106 separated events into different settings and by the duration of the event (i.e., exposure window,
107 in days) reported in each paper. Twelve event types were chosen to classify each event/setting
108 described in a paper (Table 2). To describe individual contributions to transmission, we extracted
109 data on index cases for whom contacts were followed to identify secondary transmission. We
110 only entered data from papers where it was clear from the methods that contact tracing was done
111 for at least one week to capture secondary transmission from individual index cases. For studies

112 that did not report SARS-CoV-2 variants, we imputed the dominant variant from CoVariants
113 data for the country and time period of interest [23]. See the Supplementary Material for
114 additional details on the identification of papers, data extraction (Supplementary Tables S2–S3),
115 and assessment of study bias.

116

117 **Statistical analyses**

118 To characterize the type and quality of information that we were able to extract about
119 transmission events, we performed a descriptive analysis of event data including the number of
120 each chosen event type, starting year of the data, focal countries, diagnostic methods, event
121 duration, and level of missingness for all variables. Because not all individuals potentially
122 exposed during an event were tested in each study, secondary attack rates for individual events
123 were calculated separately using the total number of exposed individuals or the total number
124 tested. If either of these quantities were missing, the value was imputed based on the value
125 present (i.e., assuming the number tested was equal to the number exposed or vice versa).
126 Sensitivity of results to this choice of denominator was assessed in the meta-analysis of events
127 (see Supplementary Material).

128 To describe the amount of variation in attack rates across studies and events and to
129 identify which settings had the highest SARS-CoV-2 attack rates, a meta-analysis was performed
130 on secondary attack rates across event types using the *metafor* package in R v4.2.2 [24]. We
131 converted secondary attack rates for each event to Freeman-Tukey double arcsine transformed
132 proportions [25] and calculated the sampling variance. We fit a hierarchical model with a nested
133 random effect for event within study and no fixed effects to assess the heterogeneity in
134 secondary attack rates attributable to these factors using restricted maximum likelihood. We

135 calculated I^2 , the percentage of variance attributable to true heterogeneity, for each random effect
136 [26] and used Cochran's Q test to test if estimated heterogeneity in secondary attack rates was
137 greater than expected from the sampling error alone. We then fit additional mixed-effects models
138 that included the same random effects but also event type and event duration as fixed effects.
139 Cochran's Q was performed on these model to assess whether residual heterogeneity in
140 secondary attack rates was greater than expected after accounting for sampling error and fixed
141 effects. Fitted coefficients and 95% confidence intervals (CI) from meta-analysis were back-
142 transformed to proportions using the geometric mean of the tested individuals across all studies
143 in each event type [25]. These back-transformed proportions are referred to as "meta-analysis
144 estimated secondary attack rates" or "meta-analysis estimated mean attack rates" in the text and
145 figures. For comparison with meta-analysis estimates, we also calculated the median secondary
146 attack rate and interquartile range across events for each chosen event type.

147 To characterize the individual offspring distribution for SARS-CoV-2, the overall
148 distribution of secondary cases generated by each identified index cases was fit to a negative
149 binomial distribution, following Lloyd-Smith et al. [11]. We estimated the percentile of index
150 cases producing 80% of all secondary infections using a formula and code from Endo et al. [27].

151 Our last aim for the study was to identify recognizable characteristics of superspreading
152 individuals. Based on the availability of demographic characteristics and other features of index
153 cases in the literature, we examined differences in distributions of secondary cases produced by
154 index cases according to sex, presence/absence of symptoms, age, real-time PCR cycle threshold
155 (Ct) value, and the total number of contacts each index case had. Additional statistical tests
156 compared these listed factors between "superspreaders" (index cases with >5 secondary cases,
157 following Adam et al. [3]) and "non-superspreaders" (index cases with ≤ 5 secondary cases): Chi-

158 square tests of proportions to compare the proportion of women, the proportion of symptomatic
159 cases, and proportion of adults or across age bins; Student's t-tests to compare mean age and Ct
160 value; and a Kruskal-Wallis test to compare the highly skewed distributions of total contacts
161 among index cases. All statistical tests used $\alpha = 0.05$ as the statistical significance threshold to
162 identify whether superspreaders were overrepresented among certain demographic groups.

163

164 **RESULTS**

165 **Study selection**

166 We identified 13,632 articles from the four databases searched, representing 8,339 unique
167 references (Figure 1). Of these, we excluded 7,358 records during the abstract review. For the
168 981 records that underwent full text review, we excluded 384 records that were reviews or letters
169 to the editor without data, contained no data on our variables of interest, or were duplicate
170 records (preprints, true duplicates, or duplicated datasets). A total of 598 papers were assessed
171 for eligibility for data extraction and a further 107 papers were excluded that did not contain
172 sufficient data on our outcome variables of interest or were duplicates (Figure 1). We extracted
173 data from 491 studies: 232 studies provided event data only, 195 studies provided individual
174 index case data only, and 64 studies provided both data types, yielding evidence from 592
175 distinct events and 9,883 index cases. The 491 analyzed studies were from 67 countries, with
176 most from China (26%), the USA (17%), and South Korea (5%) (Supplementary Figure S1A).
177 Although our search included two-thirds of 2021, nearly all studies covered data from 2020
178 (94% of events, 99% of index case symptom onset or positive test dates).

179

180 **Characteristics of events**

181 Descriptive analyses were used to characterize the type and quality of information about
182 transmission events present in the literature. Event data were most commonly from the USA
183 (27%), China (15%), the UK (8%), and South Korea (6%) (Supplementary Figure S1B).
184 Published papers were missing information on many variables that we aimed to extract about
185 events (Supplementary Figure S2A). Of the 46 target data fields from articles about events, 17
186 had high data completeness (>80%), including those for the study and event metadata, event
187 description, time period of the event (describing the start and end dates of exposure), location of
188 the event (country and state/province or city), and number of exposed individuals and secondary
189 cases (Supplementary Table S3). Event durations were highly skewed, with a median duration of
190 34 days and an interquartile range of 13–60 days (Supplementary Figure S3). Studies used a
191 variety of diagnostic methods to identify SARS-CoV-2 cases, though PCR was the dominant
192 method (Supplementary Figure S4A). Other approaches included antigen tests, retrospective case
193 identification by serology, diagnosis via symptoms or chest tomography in early papers, or a
194 mixture of approaches. Because most studies covered events prior to emergence of variants, most
195 events (N = 532, 90%) likely involved only wild-type/ancestral SARS-CoV-2, while 14 events
196 involved Alpha, six Beta, eight Delta, and 31 likely included a mixture of variants (e.g., during
197 periods of variant emergence and replacement of the dominant variant).

198

199 **Heterogeneity in event secondary attack rates**

200 Meta-analysis of secondary attack rates was performed to describe variation in attack
201 rates across studies and events and to identify which settings had the highest attack rates.
202 Secondary attack rates varied substantially within and among event types (Figure 2).
203 Interquartile ranges of attack rates were lower for transport (0–11%), hospital/healthcare (1–

204 20%), and mixed events (3–12%), whereas congregate housing (9–63%), households (15–60%),
205 social venues (8–53%), and cruise ships (9–41%) had higher heterogeneity, with some events
206 reporting attack rates of 100% (Table 2). Meta-analysis of secondary attack rates including a
207 nested random effect for event within study detected significant heterogeneity in secondary
208 attack rates ($I^2 = 99\%$, Cochran's $Q_{E,591} = 141,765$, $P < 0.0001$). The random effect for study
209 accounted for most of the heterogeneity ($I^2_{study} = 58\%$), followed by event nested within study
210 ($I^2_{event} = 41\%$). Addition of a fixed effect for event type to the model indicated that secondary
211 attack rates varied significantly across event types (Cochran's $Q_{M,11} = 122$, $P < 0.0001$). Meta-
212 analysis estimated mean attack rates were lowest for shopping (0%), hospitals and healthcare
213 (6%), transportation other than cruise ships (9%), and schools (11%) (Figure 2). Comparatively,
214 estimated mean attack rates were two to three times higher (25–35%) in nursing homes, cruise
215 ships, households, and other congregate housing settings (e.g., homeless shelters, prisons).
216 Models including event duration and an interaction term between event type and event duration
217 as additional fixed effects found similar levels of heterogeneity (Cochran's $Q_{M,23} = 135$, $P <$
218 0.0001) and identified a common trend of decreasing attack rates with longer event durations
219 across different event types, with the exception of cruise ships and shopping (Supplementary
220 Figure S5).

221

222 **Characteristics of individual index cases**

223 Descriptive analyses were also used to characterize the type and quality of information
224 about individual index cases found in published studies. Individual index case data with
225 offspring distributions overwhelmingly came from China (36%) and India (35%)
226 (Supplementary Figure S1C). Index case data exhibited higher missingness compared to events

227 (Supplementary Figure S2B): of the 74 data fields that we extracted for individual index cases,
228 the highest completeness (>60%) was seen for study and index case numbers, location of the
229 index case (country and state/province or city), total number of contacts infected, method of
230 testing for the index case and contacts, and SARS-CoV-2 variant (Supplementary Table S4). We
231 identified five key characteristics of index cases that could be related to superspreading, though
232 most of these were also missing from the published literature: 46% of cases included data on age,
233 48% on sex, 10% on presence/absence of symptoms, 6% on total number of contacts, and only
234 2% had Ct values reported. A total of 5,437 index cases (55%) contained data on at least one of
235 these five variables. Diagnostic methods for identification of individual index cases and their
236 associated secondary cases were only reported in 61% of cases, with PCR as the primary
237 approach (Supplementary Figure S4B,C). The majority of index cases (N = 8,565, 87%) were
238 assumed to be infected with wild-type SARS-CoV-2 based on location and timing of the study or
239 test confirmation date. A mixture of variants was likely in 1,282 cases (13%), while one index
240 case was reported with Alpha, two Beta, 11 Delta, and 22 Epsilon.

241

242 **Heterogeneity in transmission across individual index cases**

243 A third goal of this analysis was to describe the individual offspring distribution for
244 SARS-CoV-2 based on reported index cases. Most index cases (67%) did not transmit SARS-
245 CoV-2 to another person and 17% transmitted to only one other individual (Figure 3). There
246 were 287 “superspreaders” with >5 contacts infected, representing 3% of index cases from the
247 included studies. The distribution of secondary infections fit a negative binomial distribution
248 with a mean of 0.88 (CI: 0.84–0.92) and a dispersion parameter k of 0.27 (CI: 0.25–0.28). Using
249 the formula from Endo et al. [27] and the estimated mean and k for the negative binomial

250 distribution, the top 17% most infectious index cases would be expected to generate 80% of all
251 secondary cases.

252

253 **Qualities of superspreaders**

254 Finally, our analysis sought to identify qualities of index cases that were associated with
255 being a superspreader (index cases with >5 secondary cases) compared to non-superspreaders
256 (Table 3). The proportion of index cases with reported symptoms was higher in superspreaders
257 (89%) than non-superspreaders (76%; $\chi^2_1 = 5.4$, $P = 0.02$). Superspreaders had more than two
258 times the mean number of contacts (79) compared to non-superspreaders (36; $\chi^2_1 = 56.6$, $P <$
259 0.0001). Adults also made up a greater proportion of superspreaders (99%) than non-
260 superspreaders (84%; $\chi^2_1 = 14.1$, $P < 0.0001$). Index cases over 25 years of age were
261 overrepresented among superspreaders and no superspreaders 12 years of age and under were
262 reported (Figure 4). When age was analyzed as a continuous variable, the number of contacts
263 infected and the frequency of superspreaders increased with age, up to around 60 years of age
264 (Supplementary Figure S6). No significant differences by sex or Ct values were observed (Table
265 3). However, two adult male index cases produced the highest number of secondary infections,
266 infecting 81 of their 104 contacts and 101 of their 300 contacts, respectively. The former was a
267 lecturer in Tonghua, China [28] and the latter a fitness instructor in Hong Kong, China [29].

268 Symptomatic cases had a higher mean number of infected contacts (2.1) compared to
269 asymptomatic cases (0.7) (Table 4). The dispersion parameter k was higher for symptomatic
270 cases than asymptomatic cases (0.43 vs. 0.11), indicating lower variance in the number of
271 secondary cases produced by a symptomatic case. This variance is exemplified by the lower
272 percentage of non-transmitters (44%) and higher percentage of superspreaders (9%) among

273 symptomatic cases compared to asymptomatic cases (79% and 4%, respectively). Compared to
274 other age groups, individuals aged 49–64 years had the highest mean number of infected contacts
275 (1.2), lower variance (higher k , 0.43), and a higher percentage of superspreaders (3%). Data on
276 total reported contacts showed a different pattern, with a higher mean number of infected
277 contacts (8) as well as higher variance (lower k , 0.28) among index cases with >100 total
278 contacts compared to individuals with fewer contacts. This was accompanied by a substantially
279 higher percentage of superspreaders (28%) among individuals with >100 total contacts compared
280 to individuals with 11–100 contacts (19%) or those with 0–10 contacts (2%). Considering only
281 symptomatic adults with a known number of total contacts ($N = 129$), the percentage of
282 superspreaders was consistently smaller as the number of contacts decreased: 26% (5/19) for
283 individuals with over 100 contacts, 24% (8/34) for those with 21–100 contacts, 8% (2/24) for
284 those with 11–20 contacts, and 0% for those with 10 or fewer contacts (0/52).

285

286 **DISCUSSION**

287 In this systematic review and meta-analysis, we aimed to characterize the heterogeneity
288 in SARS-CoV-2 transmission among different settings and across individuals that has been
289 reported in published studies. Regarding transmission settings, our meta-analysis identified
290 substantial heterogeneity in attack rates across 12 chosen event types, with higher mean attack
291 rates in nursing homes, cruise ships, households, and other congregate housing settings
292 compared to shopping, hospitals and healthcare, other transportation, and schools. Regarding
293 individual transmission heterogeneity, we found that most cases did not transmit to another
294 person and that a small proportion (3%) of individuals were superspreaders (causing >5
295 secondary cases). While data on the demographics of index cases were not consistently reported

296 in the literature, the data that were available indicate that superspreaders were more likely to be
297 symptomatic than non-superspreaders, more likely to be adults (with particular
298 overrepresentation in the 49-64 age group), and had more total contacts.

299 Our ranking of event types by attack rate reinforces our existing understanding of SARS-
300 CoV-2, that transmission is more likely in dense indoor gatherings or close and frequent contact
301 among co-living individuals, especially in households [15]. Published meta-analyses covering
302 the early pandemic (pre-2021) estimated pooled household secondary attack rates of 17–21%
303 [16,18,19,30,31], with household attack rates consistently higher than those in healthcare, work,
304 or travel settings [16,19]. Our pooled household secondary attack rate over 115 events was 29%,
305 higher than these earlier studies but similar to the 31% estimate from Madewell et al. [18] for
306 studies covering July 2020 to March 2021. The higher value may be explained by the emergence
307 of the Alpha and Delta variants and the larger second and third waves of the pandemic occurring
308 in some countries during 2021.

309 The literature on SARS-CoV-2 transmission events rarely reported on the
310 epidemiological context and characteristics of different populations exposed, which could help
311 explain variation in attack rates. While the timing and location of events may help to explain
312 some of the variation within event types, the remaining variation could depend on event duration
313 (as shown by Supplementary Figure S5) and time spent indoors, types of activities occurring
314 (e.g., exercise, singing) [32,33], and the age groups present at the event. For example, the age of
315 individuals interacting in these contexts appears to also influence propensity for transmission, as
316 evidenced by the large difference in attack rates within schools versus nursing homes. Children
317 and adolescents are frequently found to have lower household infection risk than working age
318 adults [18,19,21,31] and older adults have higher risk of infection and severe disease than

319 younger ages [18,31]. In studies that assessed transmission among school-aged children,
320 teachers, and their household contacts, attack rates among children at school were lower than
321 among teachers and the household contacts of children and teachers [34,35]. Variation in the
322 stringency of interventions (e.g. masking requirements, physical distancing, lockdowns) across
323 countries and over time also could have affected attack rates across different settings. As shown
324 in Supplementary Figure S6 comparing attack rates for events in the United States and China,
325 two locations where differing stringency of control measures were implemented, meta-analysis
326 estimated attack rates were lower across event types for China, though the largest differences
327 between countries were observed for transmission in social venues and mixed settings.
328 Environmental factors such as humidity, room size, ventilation, and air flow [5] could also
329 augment transmission across settings but these were very rarely reported in the literature.

330 Analysis of index case demographics also highlighted age as an important factor in
331 SARS-CoV-2 transmission and superspreading. While age was only reported in 46% of index
332 cases, nearly all superspreading individuals were adults and there were no reported
333 superspreaders 12 years of age and under, which is consistent with other reviews of SARS-CoV-
334 2 superspreading [36]. Individual and age-related heterogeneity in the amount and assortative
335 patterns of social contacts likely influence superspreading as well. Evidence supports lower
336 transmission from children compared to adults, but effect sizes have been small in some studies
337 [16,21,30,37]. Remaining heterogeneity in individual infectiousness may derive from differences
338 in genetic susceptibility [38,39], body size (accounting for age) [40], baseline lung volume and
339 function [41], immunocompromising disease or co-infection [42,43], or the loudness and wetness
340 of speech [32]. The relative importance of these characteristics to SARS-CoV-2 transmission at a
341 population level are unknown and may be challenging to measure and report at scale. Future

342 work on COVID-19 and other respiratory diseases should address these hypotheses.

343 Our results indicate substantial heterogeneity in transmission from individuals, observed
344 in other studies [37,41,44], and evidenced by the skewed degree distribution for index cases and
345 the estimate of the dispersion parameter k . Our estimate of k (0.27, CI: 0.25–0.28) is within the
346 range of previous estimates for a similar period of the pandemic, with values frequently in the
347 range of 0.1–0.7 [3,7,8,27,45]. Caution should be taken when interpreting k values, which are
348 sensitive to changes in the tails of a distribution, such as superspreaders or individuals that cause
349 no secondary infections. Without robust isolated case finding and follow-up, contact tracing
350 efforts may undercount the number of zeroes, biasing k upwards [46,47]. Alternatively,
351 backwards contact tracing may be susceptible to attachment bias, where infections are
352 preferentially attributed to a known superspreader rather than a separate (known or unknown)
353 transmitter [47]. Additionally, there may be publication bias or more complete contact tracing for
354 large outbreaks with an individual superspreader or with high attack rates [15,47]. These effects
355 would bias k downwards and inflate meta-analysis estimated attack rates across event types. It
356 may also lead to the overestimation of the proportion of index cases that are superspreaders.
357 Without knowledge of the relative impact of these biases, it is challenging to interpret whether k
358 is a true representation of SARS-CoV-2 transmission heterogeneity. To improve inference on
359 individual heterogeneity of transmission from outbreak investigations, we recommend that
360 contact tracing efforts use both backward and forward contact tracing [15,21,48], with sufficient
361 follow-up time to identify non-infecting individuals, and complete reporting of contact tracing
362 efforts (e.g., anonymized line lists with infector-infectee and other demographic information).

363 While our systematic review is the most comprehensive assessment of SARS-CoV-2
364 superspreading to date, a principal limitation of our analysis was the incomplete data available in

365 the published literature. Beyond information provided about the timing and location of events,
366 very few studies reported any demographics of the exposed individuals, their COVID-19
367 vaccination status (once introduced) or history of prior SARS-CoV-2 infection, or the density
368 and amount of time indoors. For individual index cases, some studies reported demographic
369 information and the presence/absence of symptoms, but this atypical. We also experienced
370 difficulty with deducing whether contact tracing was performed for all reported cases in
371 transmission chains, especially for terminal nodes. It was not always clear whether cases did not
372 transmit or whether data were missing due to lack of contact tracing, so these cases had to be
373 omitted from the analysis. Testing and tracing policies likely differed between countries, which
374 would affect the collection of index cases that ended up in our review. For this reason, data on
375 index cases are missing from many countries and transmission chains from some countries may
376 be less complete than others. Similarly, the effectiveness of testing and tracing policies varies
377 across settings (e.g., easier in households than large social gatherings), which affects the
378 completeness of transmission chains and likely which outbreaks get published. There were
379 numerous papers that we reviewed with transmission chains that were simply too incomplete or
380 uncertain for us to extract index case data from them. However, without reporting of testing and
381 tracing policies or the effectiveness of tracing efforts within each paper, or a comprehensive
382 database or systematic review of this information in the literature, these remain as uncertainties
383 that must be addressed with better data.

384 Another limitation of this review was the wide variation in case detection methods across
385 studies. Not all studies reported the total number of contacts that were tested from events and we
386 assumed in the missing cases that the number tested was the same as number exposed. Our
387 sensitivity analysis, using total exposed contacts for all events as the denominator for attack rates

388 instead of total tested contacts, showed that estimated mean attack rates were consistently lower
389 across event types but the ranking of event types was relatively stable (Supplementary Figure
390 S7). However, some studies reported only symptomatic cases or only performed diagnostic tests
391 (e.g., PCR) on symptomatic individuals, thereby missing all reporting of asymptomatic or
392 pausisymptomatic individuals and any secondary cases produced. These missing contacts may be
393 undercounted for both the numerator (contacts that are infected but asymptomatic) and the
394 denominator (including contacts that are asymptomatic and uninfected), which could move
395 attack rates in either direction. Limiting testing to symptomatic contacts has a more predictable
396 effect on individual case degree distributions, reducing the apparent proportion of individuals
397 that transmit and the total secondary cases among individuals that do transmit. Case
398 ascertainment also likely varied by event setting, contributing additional uncertainty in estimated
399 attack rates. For example, performing contact tracing and testing a greater number of contacts
400 was probably easier in settings with consistent or recorded populations like households, schools,
401 and nursing homes than in large social venues like nightclubs. Differences in estimated attack
402 rates by event type may be less drastic than we observed if case ascertainment could be properly
403 addressed with additional ground truth data, i.e., community asymptomatic testing.

404 Since case detection depends partly on presence of symptoms, some care should be taken
405 in interpreting the finding that superspreaders were more likely to have symptoms than non-
406 superspreaders. We performed an additional analysis on the presence of symptoms across
407 different demographic factors reported in papers (see Supplementary Table S6). The only trend
408 we saw was for age, where the presence of symptoms was somewhat higher for older adults (49
409 and older). This may have slightly skewed detection of superspreaders among older adults.
410 However, there were still hundreds of children with symptoms reviewed as index cases, so there

411 were ample opportunities for them to be identified as superspreaders. Therefore, we remain
412 confident in our findings about the rarity of superspreaders among children. However, data from
413 human challenge trials with SARS-CoV-2 have shown that individuals with the highest viral
414 emissions did not have the most severe symptoms, but these super-emitters were also not
415 asymptomatic [41]. These super-emitters, and the majority of superspreaders reported in the
416 literature, tend to have mild to moderate symptoms [36,41]. While the importance of
417 asymptomatic transmission of SARS-CoV-2 should be acknowledged, numerous studies have
418 shown that transmission is more likely from symptomatic individuals compared to completely
419 asymptomatic individuals [18,21,49–51]. However, additional studies that overcome issues of
420 case ascertainment should be done to assess the role of asymptomatic individuals in SARS-CoV-
421 2 superspreading.

422 To improve the field and our understanding of the drivers of heterogeneity in
423 transmission, we propose standard and consistent reporting on transmission for all outbreaks, as
424 feasible, including details on the epidemiological context of transmission events and complete
425 line lists of cases following contact tracing, with information on case demographics (age, sex,
426 occupation), diagnosis (presence/absence of symptoms, symptom description, test date and
427 results), the duration of contact tracing, and the total number of contacts and the demographic
428 information for contacts (see Appendix 2). Details on the duration of contact tracing should
429 include the entire time period of case finding and how long cases were followed to detect any
430 secondary cases. We recognize the challenge of collecting, storing, and sharing identifiable data
431 from outbreak investigations while continuing to assure confidentiality and improve trust in the
432 health system. However, developing such a reporting system should be a priority for public
433 health as the information has important implications for reducing the spread of infectious

434 pathogens.

435 Our comprehensive review found substantial heterogeneity in the transmission of SARS-
436 CoV-2, highlighting the settings and individual characteristics that might be most important to
437 target for controlling superspreading. Secondary attack rates were highest in co-living situations
438 where prolonged contact between individuals facilitated transmission, though there was
439 substantial variation in attack rates within similar settings that remained unexplained and could
440 be disentangled in future meta-analyses focused on the relative influence of built environment,
441 social setting, and control measures on transmission. Given the moderate attack rates among
442 minors in school and the rarity of children among superspreaders, interventions targeting these
443 age groups may be less efficient at preventing SARS-CoV-2 superspreading and could be
444 deprioritized in favor of interventions focusing on adults [21,52], especially those with
445 symptoms and individuals with many daily close contacts. Acknowledging that there remain
446 substantial gaps in data that limit our inference about superspreading, we advocate for consistent
447 reporting on infectious disease outbreaks, ideally with detailed line lists, to facilitate knowledge
448 synthesis about transmission patterns and superspreading in the future. Our review only covered
449 the first phase of the pandemic, so important questions remain about whether patterns in attack
450 rates and individual-level transmission still apply to later pandemic phases with significant
451 population-level immunity. Enhanced reporting of outbreak data would expedite such future
452 investigations.

453

454 **DATA AVAILABILITY**

455 All the data were from publicly available databases. The complete database of extracted
456 information from included studies is provided in Appendix 3.

457

458 **FINANCIAL SUPPORT**

459 This research was funded by the World Health Organization. The topic of the review was
460 proposed by the senior author (ESG) and the WHO co-author (MVK) reviewed the manuscript
461 and agreed to publish as a co-author; the funding agency had no role in study design, data
462 collection and analysis, or decision to publish.

463

464 **COMPETING INTERESTS**

465 The authors declare none.

466

467 **REFERENCES**

- 468 1. **Davis JT, et al.** Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature*
469 2021; **600**: 127–132.
- 470 2. **Lau MSY, et al.** Characterizing superspreading events and age-specific infectiousness of
471 SARS-CoV-2 transmission in Georgia, USA. *Proceedings of the National Academy of*
472 *Sciences* 2020; **117**: 22430–22435.
- 473 3. **Adam DC, et al.** Clustering and superspreading potential of SARS-CoV-2 infections in Hong
474 Kong. *Nature Medicine* 2020; **26**: 1714–1719.
- 475 4. **Liu Y, Eggo RM, Kucharski AJ.** Secondary attack rate and superspreading events for
476 SARS-CoV-2. *The Lancet* 2020; **395**: e47.
- 477 5. **Lakdawala SS, Menachery VD.** Catch me if you can: superspreading of COVID-19. *Trends*
478 *in Microbiology* 2021; **29**: 919–929.

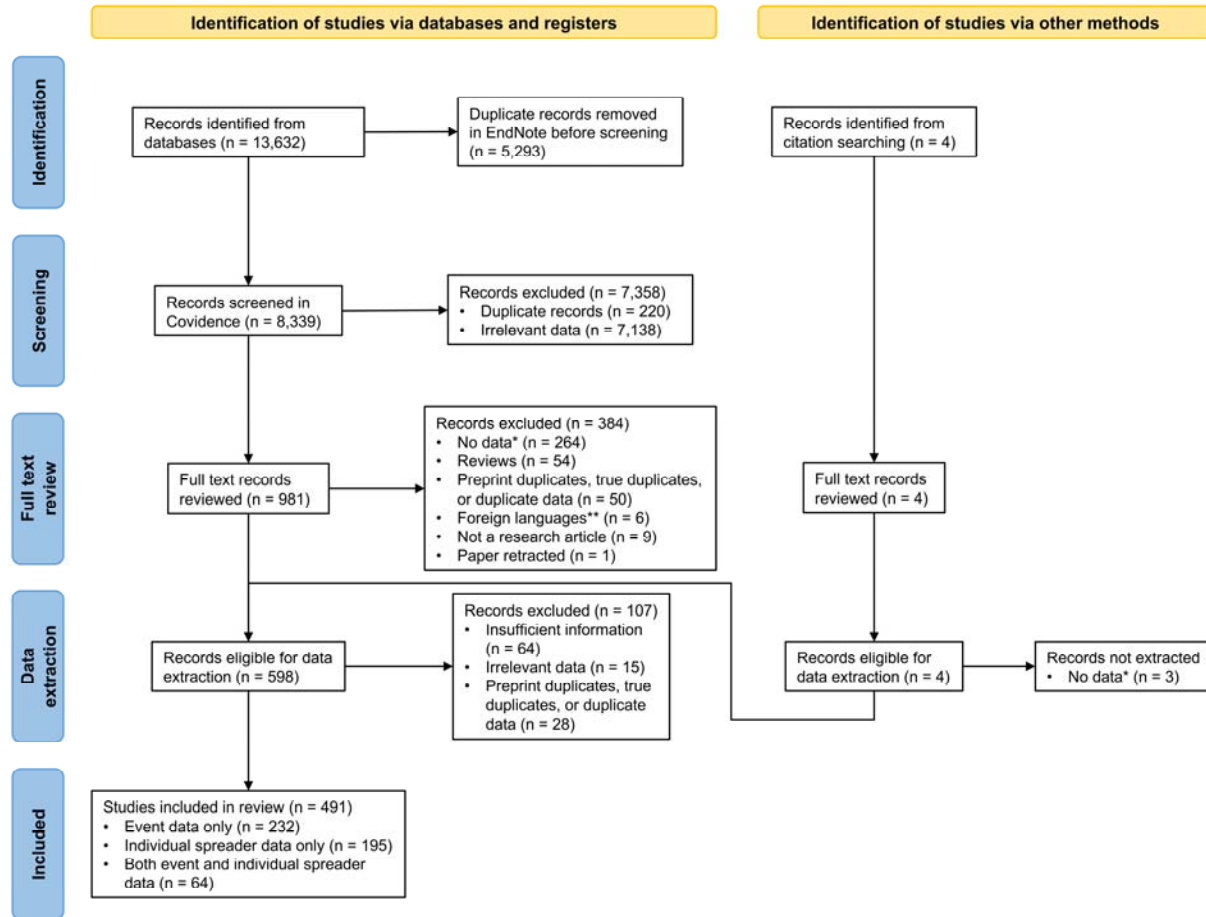
- 479 6. **Puhach O, Meyer B, Eckerle I.** SARS-CoV-2 viral load and shedding kinetics. *Nature*
480 *Reviews Microbiology* 2023; **21**: 147–161.
- 481 7. **Wang J, et al.** Superspreading and heterogeneity in transmission of SARS, MERS, and
482 COVID-19: a systematic review. *Computational and Structural Biotechnology Journal* 2021;
483 **19**: 5039–5046.
- 484 8. **Du Z, et al.** Systematic review and meta-analyses of superspreading of SARS-CoV-2
485 infections. *Transboundary and Emerging Diseases* 2022; **69**: e3007–e3014.
- 486 9. **Althaus CL.** Ebola superspreading. *The Lancet Infectious Diseases* 2015; **15**: 507–508.
- 487 10. **Paunio M, et al.** Explosive school-based measles outbreak: intense exposure may have
488 resulted in high risk, even among revaccinees. *American Journal of Epidemiology* 1998; **148**:
489 1103–1110.
- 490 11. **Lloyd-Smith JO, et al.** Superspreading and the effect of individual variation on disease
491 emergence. *Nature* 2005; **438**: 355–359.
- 492 12. **Nikolay B, et al.** Transmission of Nipah virus — 14 years of investigations in Bangladesh.
493 *The New England Journal of Medicine* 2019; **380**: 1804–1814.
- 494 13. **Lewis D.** Superspreading drives the COVID pandemic — and could help to tame it. *Nature*
495 2021; **590**: 544–546.
- 496 14. **Hamner L, et al.** High SARS-CoV-2 attack rate following exposure at a choir practice —
497 Skagit County, Washington, March 2020. *MMWR. Morbidity and Mortality Weekly Report*
498 2020; **69**: 606–610.
- 499 15. **Lee EC, et al.** The engines of SARS-CoV-2 spread. *Science* 2020; **370**: 406–407.

- 500 16. **Thompson HA, et al.** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
501 setting-specific transmission rates: a systematic review and meta-analysis. *Clinical Infectious*
502 *Diseases* 2021; **73**: e754–e764.
- 503 17. **Leclerc QJ, et al.** What settings have been linked to SARS-CoV-2 transmission clusters?
504 *Wellcome Open Research* 2020; **5**: 83.
- 505 18. **Madewell ZJ, et al.** Factors associated with household transmission of SARS-CoV-2: an
506 updated systematic review and meta-analysis. *JAMA Network Open* 2021; **4**: e2122240.
- 507 19. **Koh WC, et al.** What do we know about SARS-CoV-2 transmission? A systematic review
508 and meta-analysis of the secondary attack rate and associated risk factors. *PLOS ONE* 2020;
509 **15**: e0240205.
- 510 20. **Ramírez-del Real T, et al.** Individual factors associated with COVID-19 infection: a
511 machine learning study. *Frontiers in Public Health* 2022; **10**: 912099.
- 512 21. **Gupta M, et al.** Contact tracing of COVID-19 in Karnataka, India: superspreading and
513 determinants of infectiousness and symptomatic infection. *PLOS ONE* 2022; **17**: e0270789.
- 514 22. **Page MJ, et al.** The PRISMA 2020 statement: an updated guideline for reporting systematic
515 reviews. *Systematic Reviews* 2021; **10**: 89.
- 516 23. **Hodcroft EB.** *CoVariants: SARS-CoV-2 Mutations and Variants of Interest.*
517 2021(<https://covariants.org/>).
- 518 24. **Viechtbauer W.** Conducting meta-analyses in R with the metafor package. *Journal of*
519 *Statistical Software* 2010; **36**: 1–48.
- 520 25. **Lin L, Xu C.** Arcsine-based transformations for meta-analysis of proportions: pros, cons,
521 and alternatives. *Health Science Reports* 2020; **3**: e178.

- 522 26. **Senior AM, et al.** Heterogeneity in ecological and evolutionary meta-analyses: its magnitude
523 and implications. *Ecology* 2016; **97**: 3293–3299.
- 524 27. **Endo A, et al.** Estimating the overdispersion in COVID-19 transmission using outbreak sizes
525 outside China. *Wellcome Open Research* 2020; **5**: 67.
- 526 28. **Yao L, et al.** COVID-19 super spreading event amongst elderly individuals — Jilin
527 Province, China, January 2021. *China CDC Weekly* 2021; **3**: 211–213.
- 528 29. **Chu DKW, et al.** SARS-CoV-2 superspread in fitness center, Hong Kong, China, March
529 2021. *Emerging Infectious Diseases* 2021; **27**: 2230–2232.
- 530 30. **Madewell ZJ, et al.** Household transmission of SARS-CoV-2: a systematic review and
531 meta-analysis. *JAMA Network Open* 2020; **3**: e2031756.
- 532 31. **Bi Q, et al.** Insights into household transmission of SARS-CoV-2 from a population-based
533 serological survey. *Nature Communications* 2021; **12**: 3643.
- 534 32. **Asadi S, et al.** Aerosol emission and superemission during human speech increase with voice
535 loudness. *Scientific Reports* 2019; **9**: 2348.
- 536 33. **Mürbe D, et al.** Aerosol emission in professional singing of classical music. *Scientific*
537 *Reports* 2021; **11**: 14861.
- 538 34. **Aizawa Y, et al.** Coronavirus disease 2019 cluster originating in a primary school teachers’
539 room in Japan. *The Pediatric Infectious Disease Journal* 2021; **40**: e418–e423.
- 540 35. **Gettings JR, et al.** SARS-CoV-2 transmission in a Georgia school district — United States,
541 December 2020–January 2021. *Clinical Infectious Diseases* 2022; **74**: 319–326.
- 542 36. **Brainard J, et al.** Super-spreaders of novel coronaviruses that cause SARS, MERS and
543 COVID-19: a systematic review. *Annals of Epidemiology* 2023; **82**: 66-76.e6.

- 544 37. **Ke R, et al.** Daily longitudinal sampling of SARS-CoV-2 infection reveals substantial
545 heterogeneity in infectiousness. *Nature Microbiology* 2022; **7**: 640–652.
- 546 38. **Severe Covid-19 GWAS Group, et al.** Genomewide association study of severe Covid-19
547 with respiratory failure. *The New England Journal of Medicine* 2020; **383**: 1522–1534.
- 548 39. **Zhao J, et al.** Relationship between the ABO blood group and the coronavirus disease 2019
549 (COVID-19) susceptibility. *Clinical Infectious Diseases* 2021; **73**: 328–331.
- 550 40. **Edwards DA, et al.** Exhaled aerosol increases with COVID-19 infection, age, and obesity.
551 *Proceedings of the National Academy of Sciences* 2021; **118**: e2021830118.
- 552 41. **Zhou J, et al.** Viral emissions into the air and environment after SARS-CoV-2 human
553 challenge: a phase 1, open label, first-in-human study. *The Lancet Microbe* 2023; **4**: e579–
554 e590.
- 555 42. **Caillard S, et al.** SARS-CoV-2 viral dynamics in immunocompromised patients. *American*
556 *Journal of Transplantation* 2021; **21**: 1667–1669.
- 557 43. **Aytillo T, et al.** Shedding of viable SARS-CoV-2 after immunosuppressive therapy for
558 cancer. *New England Journal of Medicine* 2020; **383**: 2586–2588.
- 559 44. **Chen PZ, et al.** Heterogeneity in transmissibility and shedding SARS-CoV-2 via droplets
560 and aerosols. *eLife* 2021; **10**: e65774.
- 561 45. **Wegehaupt O, Endo A, Vassall A.** Superspreading, overdispersion and their implications in
562 the SARS-CoV-2 (COVID-19) pandemic: a systematic review and meta-analysis of the
563 literature. *BMC Public Health* 2023; **23**: 1003.
- 564 46. **Zhao S, et al.** Inferencing superspreading potential using zero-truncated negative binomial
565 model: exemplification with COVID-19. *BMC Medical Research Methodology* 2021; **21**: 30.

- 566 47. **Taube JC, Miller PB, Drake JM.** An open-access database of infectious disease
567 transmission trees to explore superspreader epidemiology. *PLOS Biology* 2022; **20**:
568 e3001685.
- 569 48. **Endo A, et al.** Implication of backward contact tracing in the presence of overdispersed
570 transmission in COVID-19 outbreaks. *Wellcome Open Research* 2021; **5**: 239.
- 571 49. **Qiu X, et al.** The role of asymptomatic and pre-symptomatic infection in SARS-CoV-2
572 transmission—a living systematic review. *Clinical Microbiology and Infection* 2021; **27**:
573 511–519.
- 574 50. **Buitrago-Garcia D, et al.** Occurrence and transmission potential of asymptomatic and
575 presymptomatic SARS-CoV-2 infections: update of a living systematic review and meta-
576 analysis. *PLOS Medicine* 2022; **19**: e1003987.
- 577 51. **James A, et al.** Model-free estimation of COVID-19 transmission dynamics from a complete
578 outbreak. *PLOS ONE* 2021; **16**: e0238800.
- 579 52. **Davies NG, et al.** Age-dependent effects in the transmission and control of COVID-19
580 epidemics. *Nature Medicine* 2020; **26**: 1205–1211.
- 581



582

583 **Figure 1.** PRISMA flow diagram for the systematic review and meta-analysis of SARS-CoV-2

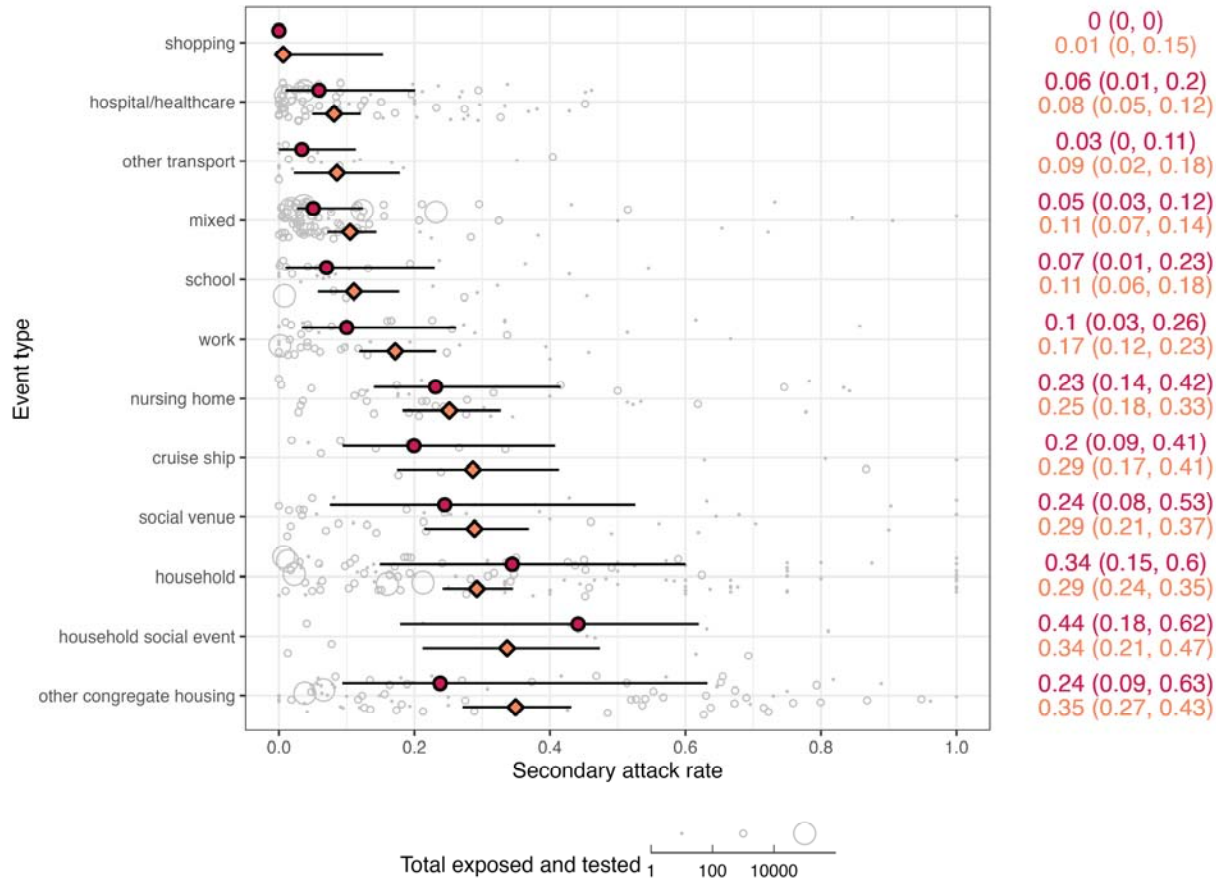
584 superspreading reported in the published literature. *There were 4 types of data that we sought to

585 include: 1) transmission chain; 2) number of index cases, number of contacts, and number of

586 infected contacts; 3) number of index cases and number of infected contacts; or 4) secondary

587 attack rate. **Languages other than Spanish, Chinese, French, Turkish, German, and Portuguese.

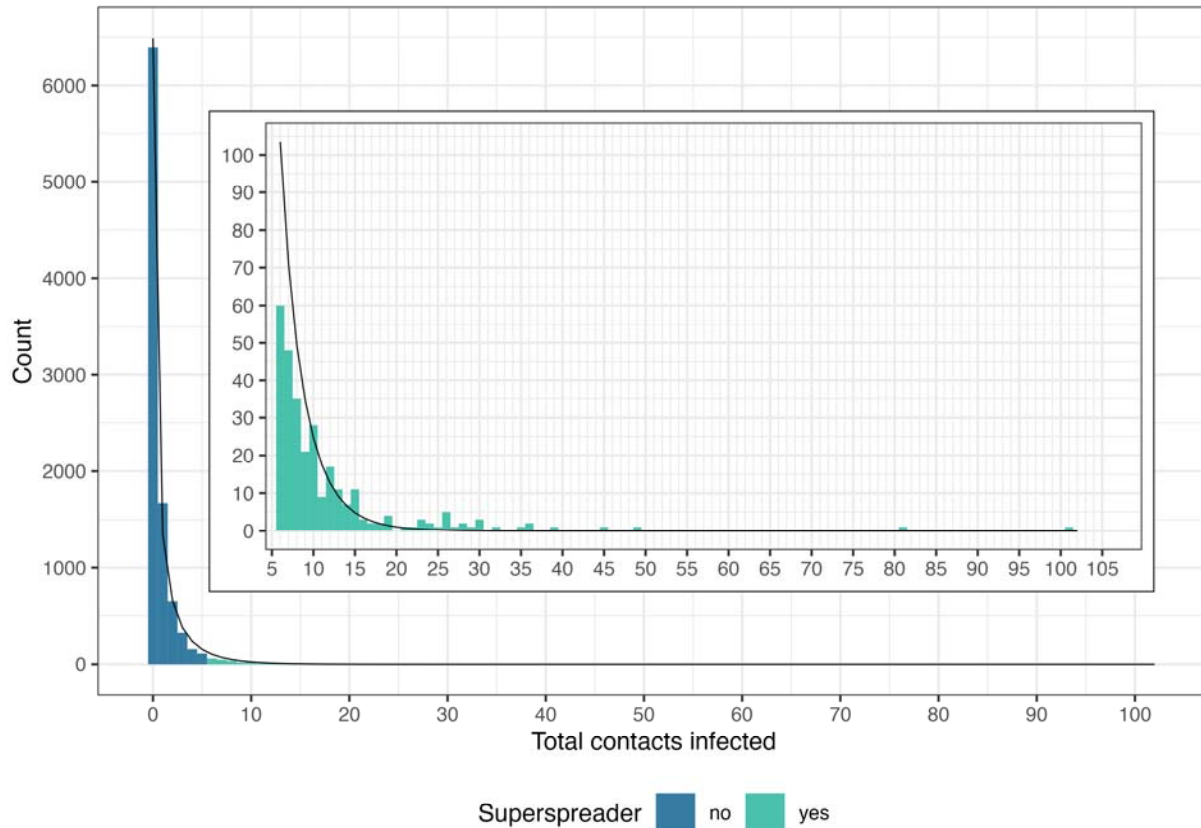
588



589

590 **Figure 2.** SARS-CoV-2 secondary attack rates across 12 event types occurring between
 591 December 2019 and August 2021 reported in the literature across 592 events from 296 studies.
 592 Individual event data secondary attack rates are shown as grey bubbles, varying in size according
 593 to the total number of individuals exposed and tested from the event. Median secondary attack
 594 rate for each event type is shown as red circle with a line representing the interquartile range;
 595 values are in red on the right side of the figure. Meta-analysis estimated secondary attack rate for
 596 each event type is shown as an orange diamond with a line representing the estimated 95%
 597 confidence interval; values are in orange on the right side of the figure. Event types were ranked
 598 by increasing estimated mean secondary attack rate along the left axis.

599



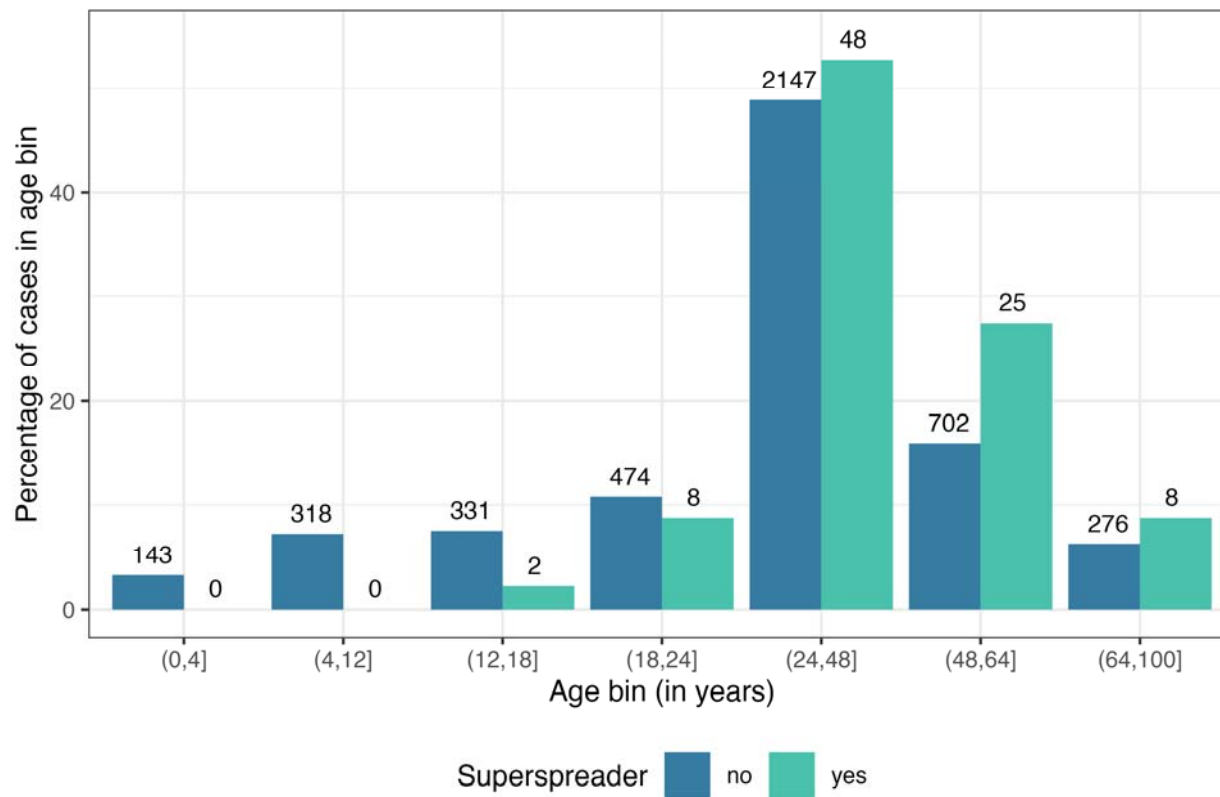
600

601 **Figure 3.** Distribution of secondary contacts infected by individual index cases (N = 9,591) for
602 SARS-CoV-2 cases occurring between December 2019 and July 2021 reported in 259 studies.

603 The black line shows the fit of the distribution to the expected negative binomial distribution.

604 The inset shows a portion of the same data to highlight the distribution of superspreaders (index
605 cases with >5 secondary cases).

606



607

608 **Figure 4.** Comparison of the age distribution of superspreading index cases. The bars show the
609 percentage of individuals within an age bin across superspreaders (index cases with >5
610 secondary cases) and non-superspreaders. Numbers above the bars display the raw totals and
611 percentages are shown in Table 3.

612

613 **Table 1.** Description of variables extracted from papers in the systematic review of SARS-CoV-
 614 2 superspreading from December 2019 to July 2021.

All papers	<ul style="list-style-type: none"> • Title • Author(s) • Publication year, volume, and issue • Journal • Study location(s) <ul style="list-style-type: none"> ○ Country ○ Administrative unit(s): state/province, county, city • General study time period (e.g., start and end year/month of data collection) • Diagnostic testing method (PCR, serology, rapid antigen tests, symptom diagnosis, or mixed) • Variant name • Any reported prevention measures implemented in the event (e.g., masking, social distancing) • Number of exposed people with reported demographic characteristics (age, sex) and vaccination status
Events	<ul style="list-style-type: none"> • Type of event/setting (e.g. nursing home residents, household transmission study, or school) • Start and end date of event
Index case	<ul style="list-style-type: none"> • Demographic characteristics (age, sex, occupation) <ul style="list-style-type: none"> ○ Age group: ≤ 4 years, 5–12 years, 13–18 years, 19–24 years, 25–48 years, 49–64 years, ≥ 65 years • Symptom onset (if applicable) and diagnosis dates • Symptoms (text descriptions or presence/absence) • Real-time PCR cycle threshold (Ct) value

	<ul style="list-style-type: none">• Specimen type• Clinical outcome (if applicable)• Setting of contact (e.g., work, social, and school)
--	--

615 **Table 2.** Types of SARS-CoV-2 secondary transmission events occurring between December
 616 2019 and August 2021 reported in the literature. Heterogeneity across event types was assessed
 617 based on the variance and interquartile range of secondary attack rates. Outlier events were
 618 identified for each event type as events that exceeded the estimated upper confidence interval of
 619 the meta-analysis estimated SAR for that event type or were greater than 50%.

Event type	Description of outbreak location	N	Minimum and maximum secondary attack rate
1	cruise ship or other densely populated watercraft (e.g., fishing vessel, aircraft carrier)	16	0.02, 1
2	transport mode other than ships (e.g., airplane, train, car)	20	0, 0.4
3	households, defined as co-living individuals or close contacts who always meet each other but possibly not living together (e.g., couples in romantic relationship)	115	0, 1
4	hospital or healthcare facility, including patients, healthcare workers, and nursing home workers (if worker data was provided separately from nursing home residents)	89	0, 0.46
5	workplace (e.g., office), including correctional officers and teachers and staff at schools	51	0, 0.86
6	school (data on students only)	32	0, 0.54
7	public social venue (e.g., bar, concert, sporting event)	39	0, 1
8	private social event with members of multiple households (e.g., dinner with neighbors or extended family)	12	0.01, 0.81
9	shopping (activities in shops, markets, and department stores)	2	0, 0
10	nursing home or long-term care facility (residents only or residents and healthcare workers if not described separately in the paper)	41	0, 0.84

11	congregate housing other than nursing home or long-term care facility (e.g., homeless shelter, prison, summer camp)	84	0, 1
12	mixed locations, included any combination of the above but not described separately in the paper	91	0, 1

620 **Table 3.** Statistical comparisons of SARS-CoV-2 superspreaders to non-superspreaders based on
 621 features reported in the literature in 259 studies for cases occurring between December 2019 and
 622 July 2021.

Feature of comparison	Percentage or estimated mean for non-superspreaders (total observations)	Percentage or estimated mean for superspreaders (total observations)	Statistical test results
Female	40% (N = 4,543)	38% (N = 102)	$\chi^2_1 = 0.09, P = 0.76$
Presence of symptoms (symptomatic)	76% (N = 841)	89% (N = 70)	$\chi^2_1 = 5.4, P = 0.02$
Age (in bins)	(N = 4,391)	(N = 91)	$\chi^2_6 = 21.7, P = 0.001$
≤4 years	3%	0%	
5–12 years	7%	0%	
13–18 years	8%	2%	
19–24 years	11%	9%	
25–48 years	49%	53%	
49–64 years	16%	27%	
≥65 years	6%	9%	
Age (≥18 years)	84% (N = 4,391)	99% (N = 91)	$\chi^2_1 = 14.1, P < 0.0001$
Age (in years)	34.8 (N = 4,391)	43.8 (N = 91)	$t_{94.4} = 5.2, P < 0.0001$
Ct value	26.7 (N = 140)	24.8 (N = 10)	$t_{10.1} = -0.8, P = 0.45$
Total contacts	36 (N = 471)	79 (N = 59)	$\chi^2_1 = 56.6, P < 0.0001$

623 **Table 4.** Summary statistics describing the distribution of secondary cases among individual SARS-CoV-2 index cases occurring
 624 between December 2019 and July 2021 reported in the literature across 259 studies.

Data	Sample size	Percentage with 0 contacts infected	Percentage with 1–5 contacts infected	Percentage with >5 contacts infected	Maximum contacts infected	Estimated mean contacts infected (95% CI)	Estimated overdispersion, k (95% CI)
All rows	9,591	67%	30%	3%	101	0.88 (0.84–0.92)	0.27 (0.25–0.28)
Female	1,866	75%	22%	2%	30	0.63 (0.56–0.71)	0.18 (0.16–0.21)
Male	2,779	74%	24%	2%	101	0.76 (0.69–0.84)	0.17 (0.15–0.19)
Asymptomatic	214	79%	17%	4%	25	0.75 (0.43–1.08)	0.11 (0.07–0.16)
Symptomatic	697	44%	47%	9%	81	2.06 (1.8–2.3)	0.43 (0.36–0.49)
Age ≤4 years	143	90%	10%	0%	3	0.2 (0.07–0.32)	0.09 (0.01–0.17)
Age 5–12 years	318	97%	3%	0%	3	0.04 (0.006–0.07)	0.03 (–0.006–0.06)
Age 13–18 years	333	94%	6%	1%	26	0.23 (0.08–0.38)	0.03 (0.01–0.05)
Age 19–24 years	482	88%	10%	2%	19	0.38 (0.23–0.52)	0.06 (0.04–0.09)
Age 25–48 years	2,195	76%	22%	2%	101	0.73 (0.65–0.82)	0.15 (0.13–0.17)
Age 49–64	727	54%	43%	3%	35	1.16 (1.01–1.31)	0.43 (0.36–0.51)
Age ≥65 years	284	58%	39%	3%	18	1 (0.79–1.21)	0.43 (0.3–0.57)

Ct value ≤ 25	67	42%	52%	6%	12	1.48 (1–1.96)	0.86 (0.31–1.42)
Ct value > 25	83	52%	41%	7%	26	1.45 (0.87–2.02)	0.4 (0.2–0.6)
0–10 total contacts	281	49%	48%	2%	9	1.12 (0.93–1.31)	0.83 (0.53–1.14)
11–100 total contacts	195	47%	34%	19%	39	3.1 (2.29–3.91)	0.32 (0.23–0.41)
101–1000 total contacts	54	35%	37%	28%	101	8 (3.92–12.07)	0.28 (0.16–0.4)

625