

1 **Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study**

2

3 Michael Marks PhD, <sup>1,2</sup> Pere Millat-Martinez MD, <sup>3</sup> Dan Ouchi MSc, <sup>4</sup> Chrissy h. Roberts PhD, <sup>1</sup> Andrea  
4 Alemany BM, <sup>5</sup> Marc Corbacho-Monné BM, <sup>5</sup> Maria Ubals MD, <sup>6</sup> Aurelio Tobias PhD <sup>7</sup>, Cristian Tebé  
5 PhD<sup>8,9</sup>, Ester Ballana PhD<sup>4,6</sup>, Martí Vall-Mayans PhD, <sup>5,6</sup> Camila G-Beiras PhD, <sup>5</sup> Nuria Prat MSc, <sup>6</sup> Jordi  
6 Ara PhD, <sup>6</sup> Bonaventura Clotet PhD, <sup>4,5,6,10</sup> Oriol Mitjà PhD <sup>5,6,9,10,11</sup>

7

8 1. London School of Hygiene and Tropical Medicine, London, United Kingdom

9 2. Hospital for Tropical Diseases, London, United Kingdom

10 3. Barcelona Institute for Global Health – University of Barcelona, Barcelona, Spain

11 4. IrsiCaixa AIDS Research Institute, Badalona, Spain

12 5. Fight AIDS and Infectious Diseases Foundation, Badalona, Spain

13 6. Hospital Universitari Germans Trias i Pujol, Badalona Spain

14 7. Institute of Environmental Assessment and Water Research (IDAEA), Spanish Council for Scientific  
15 Research (CSIC), Barcelona, Spain

16

17 8 Biostatistics Unit, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet de  
18 Llobregat, Barcelona, Spain

19

20 9. Department of Clinical Sciences, Faculty of Medicine and Health Sciences, University of Barcelona ,  
21 Barcelona, Spain.

22

23 10. Universitat de Vic – Universtiat Central de Catalunya, Vic, Spain

24 11. Lihir Medical Centre-InternationalSOS, Lihir Island, Papua New Guinea

25

26

27 Corresponding author: Michael Marks

28 Clinical Research Department,

29 London School of Hygiene & Tropical Medicine

30 London

31 United Kingdom

32 WC1E 7HT

## 33 **ABSTRACT**

### 34 *Background*

35 There remains limited data on what variables affect risk of transmission of SARS-CoV-2 and developing  
36 symptomatic Covid-19 and in particular the relationship to viral load (VL). We analysed data from linked  
37 index cases and their contacts to explore factors associated with transmission of SARS-CoV-2.

### 38 *Methods*

39 Patients were recruited as part of a randomized control trial ,conducted between March to April 2020, that  
40 aimed to assess if hydroxychloroquine reduced transmission of SARS-CoV-2. Non-hospitalised Covid-19  
41 cases and their contacts were identified through the local surveillance system. VL, measured by  
42 quantitative PCR from a nasopharyngeal swab, was assessed at enrollment, at day 14, and whenever the  
43 participant reported Covid-19-like symptoms. Risk of transmission, developing symptomatic disease and  
44 incubation dynamics were evaluated using regression analysis.

### 45 *Findings*

46 We identified 314 cases, 282 of which had at least one contact (753 contacts in total). Ninety (33%) of  
47 282 clusters had at least one transmission event. The secondary attack rate was 16% (125/753), with a  
48 variation from 12% to 24% for VL of the index case of  $<10^6$ , and  $>10^9$  copies/mL, respectively (OR per  
49  $\log_{10}$  increase in VL 1.3 95%CI 1.1–1.6). Increased risk of transmission was also associated with  
50 household contact (OR 2.7; 1.4–5.06) and age of the contact (OR 1.02 per year; 1.01–1.04). The  
51 proportion of PCR positive contacts who developed symptomatic Covid-19 was 40.3% (181/449), with a  
52 variation from 25% to 60% for VL of the contact  $<10^7$ , and  $>10^9$  copies/mL (HR  $\log_{10}$  increase in VL  
53 1.12; 95% CI 1.05 – 1.2). Time to onset of symptomatic disease decreased from a median of 7 days (IQR  
54 5–10) for individuals with an initial viral load  $<10^7$  to 6 days (4–8) and 5 days (3–8) for individuals with  
55 an initial viral load of  $10^7$ – $10^9$  and  $>10^9$ , respectively.

### 56 *Interpretation*

57 Viral load of index cases is a leading driver of SARS-CoV-2 transmission. The risk of symptomatic  
58 Covid-19 is strongly associated with viral load of contacts at baseline and shortens the incubation time in  
59 a dose-dependent manner.

60

61 Funding: Crowdfunding campaign YoMeCorono (<https://www.yomecorono.com/>), and Generalitat de  
62 Catalunya. Support for laboratory equipment from Foundation Dormeur.

63

## 64 **Research in context**

### 65 *Evidence before this study*

66 In September 2020, we searched PubMed database for articles reporting on factors influencing  
67 transmission and the risk of developing symptomatic disease. Search terms included “Covid-19”, “SARS-  
68 CoV-2”, “transmission”, “incubation time”, and “risk”, with no language restrictions. By 20<sup>th</sup> September,  
69 various authors had reported on retrospective analyses of clusters of index cases and their corresponding  
70 contacts, as well as series of patients who developed symptomatic Covid-19 disease after PCR positive  
71 result. Besides describing the secondary attack rate, various authors identified risk factors for  
72 transmission associated with the place and duration of exposure and the lack of use of personal protective  
73 equipment. A single study suggested that symptomatic individuals might be more likely to transmit than  
74 asymptomatic cases but we found no clear evidence regarding the influence of viral load of the index case  
75 on transmission risk. Similarly, although various retrospective series of patients with positive PCR results  
76 had reported incubation times elsewhere, the characteristics of index case and contacts that may influence  
77 the risk of developing symptomatic Covid-19 and the time to this event had been barely addressed.

### 78 *Added value of this study*

79 We analyzed data from a large cluster-randomized clinical trial on post-exposure therapy for Covid-19  
80 that provide new information on SARS-CoV-2 transmission dynamics. Several design components add  
81 value to this dataset. Notably, quantitative PCR was available for the index cases to estimate risk of  
82 transmission. Furthermore, quantitative PCR was also performed on asymptomatic contacts at the time of  
83 enrollment allowing to investigate the dynamics of symptomatic disease onset among them. We found  
84 that the viral load of the index case was the leading determinant of the risk of SARS-CoV-2 PCR  
85 positivity among contacts. Among contacts who were SARS-CoV-2 PCR positive at baseline, viral load  
86 significantly influenced the risk of developing the symptomatic disease in a dose-dependent manner. This  
87 influence also became apparent in the incubation time, which shortened with increasing baseline viral  
88 loads.

89 *Implication of all the available evidence*

90 Our results provide important insights into the knowledge regarding the risk of SARS-CoV-2  
91 transmission and Covid-19 development. The fact that the transmission risk is primarily driven by the  
92 viral load of index cases, more than other factors such as their symptoms or age, suggests that all cases  
93 should be considered potential transmitters irrespective of their presentation and encourages assessing  
94 viral load in cases with a larger number of close contacts. Similarly, our results regarding the risk and  
95 expected time to developing symptomatic Covid-19 encourage risk stratification of newly diagnosed  
96 SARS-CoV-2 infections based on the initial viral load.

97

98

## 99 INTRODUCTION

100 According to current evidence, Covid-19 is primarily transmitted from person to person through  
101 respiratory droplets, as well as indirect contact, through transfer of the virus from contaminated fomites to  
102 the mouth, nose, or eyes.<sup>1,2</sup> As with most respiratory viral infections there is likely to be some  
103 contribution from smaller aerosols but their relative contribution compared to droplets remains unclear.  
104 Several outbreak investigation reports have shown that Covid-19 transmission can be particularly  
105 effective in confined indoor spaces such as workplaces including factories, churches, restaurants,  
106 shopping centers, or healthcare settings.<sup>3-6</sup> In Spain, and many other countries, healthcare workers have  
107 experienced a high rate of Covid-19 infection.<sup>7</sup>

108 The availability of data regarding the factors that may enhance transmission is essential for designing  
109 interventions to control SARS-CoV-2 spread. Currently available data provide information on the risk of  
110 transmission related to the place and duration of exposure, and the use of respiratory and eye protection<sup>1,3-</sup>  
111 <sup>5,8</sup> but not on other factors related to the characteristics of index cases and their contacts. Over the course  
112 of infection, the virus has been identified in respiratory tract specimens 1–2 days before the onset of  
113 symptoms, and it can persist for prolonged periods over several weeks after the onset of symptoms in  
114 mild cases.<sup>9</sup> However, the detection of viral RNA by PCR does not necessarily equate with infectivity,  
115 and the exact relationship between viral load and risk of transmission from a case is still not clear.<sup>10,11</sup>  
116 Studies investigating case-contact pairs have reported highly variable secondary attack rates (i.e., range  
117 0.7% to 75%), depending on the type of exposure—duration, place, pre- or post-symptomatic.<sup>12-15</sup>

118 Another challenge for public health interventions is the risk stratification of infected individuals for  
119 developing symptomatic illness. On the other hand, a living systematic review estimated that the  
120 proportion of PCR-positive infected contacts that progress to symptomatic disease is approximately 70-  
121 80%.<sup>16,17</sup> Estimates of mean or median incubation period have been consistently between 5–7 days.<sup>18-20</sup>  
122 Whilst there has been a suggestion that viral load of cases may potentially be associated with risk of  
123 disease or transmission there is currently no published data directly addressing this question and little is  
124 known about factors that may contribute to variation on the risk of developing Covid-19 symptoms or the  
125 incubation periods among infected individuals.

126 The overall aim of this study was to evaluate transmission dynamics of SARS-CoV-2 in the context of a  
127 trial of post-exposure prophylaxis. Specifically, the objectives of the study were threefold: (a) to  
128 investigate the association between clinical and demographic features of cases and viral load, (b) to  
129 evaluate the effect of viral load on SARS-CoV-2 transmission to close contacts, and (c) to determine the  
130 influence of viral load in the exposed on development of symptoms and on the incubation period.

131 **METHODS**

132 *Study design*

133 This was a post-hoc analysis of data collected in the BCN PEP CoV-2 Study (NCT04304053), a cluster-  
134 randomized trial that included PCR-confirmed Covid-19 cases and their close contacts. The trial occurred  
135 between Mar 17 to Apr 28, 2020, during the SARS-CoV-2 outbreak, in three out of nine healthcare areas  
136 in Catalonia (North-East Spain): *Catalunya central, Àmbit Metropolità Nord, and Barcelona Ciutat*, total  
137 target population 4,206,440 people. The study protocol of the BCN PEP CoV-2 Study was approved by  
138 the ethics committee of Hospital Germans Trias Pujol, (Badalona, Spain). Written informed consent was  
139 obtained from all participants. Full details of the original study are reported elsewhere.<sup>21</sup>

140 Covid-19 cases were identified using the electronic registry of the Epidemiological Surveillance  
141 Emergency Service of Catalonia (SUVEC) of the Department of Health.<sup>22</sup> Following government  
142 ordinance, the SUVEC registered all new Covid-19 diagnoses occurred from March 16, 2020. The  
143 surveillance system included active tracing of all contacts with recent history of exposure, defined as  
144 being in contact with a SARS-CoV-2 PCR positive case during more than 15 minutes within two meters.

145 All Covid-19 cases included in the present analysis were non-hospitalized adults (i.e.,  $\geq 18$  years of age)  
146 with quantitative PCR result available at baseline, mild symptom onset within five days before  
147 enrollment, and no reported symptoms of SARS-CoV-2 infections in their accommodation (i.e.,  
148 household or nursing home) or workplace within the 14 days before enrollment. Contacts selected for the  
149 analysis were adults with a recent history of exposure and absence of Covid-19-like symptoms within the  
150 seven days preceding enrolment. Contacts were exposed to the index case as either a healthcare worker, a  
151 household contact, a nursing home worker, or a nursing home resident.

152 *Study procedures and data collection*

153 A dedicated outbreak field team visited cases and contacts at home or nursing home on days 1  
154 (enrollment) and 14. At the first clinical assessment on day 1 they conducted a baseline assessment,  
155 including a questionnaire for symptoms of Covid-19 and collected relevant epidemiological information  
156 using a structured interview: time of first exposure to the index case, place of contact (hospital, home,  
157 nursing care facility), routine use of a mask of both when in close proximity to the index case, the case  
158 and the contact, and sleep location concerning the index case (e.g., same room, same house). Symptoms  
159 surveillance consisted of active monitoring by phone on days 3, and 7, a home visit on day 14, and  
160 passive monitoring whenever the participants developed symptoms. Participants who developed  
161 symptoms were visited the same day they notified symptom onset (unscheduled visits) by the field team,

162 which recorded the date of symptom onset, type of symptoms from a pre-specified checklist, and  
163 symptom severity, graded on a 1-to-4 scale.

164 Serial SARS-CoV-2 PCR test and viral load titration on nasopharyngeal swab were conducted on day 1  
165 and day 14 to all participants, and on any unscheduled visit when the participant notified the onset of  
166 Covid-19 symptoms. The detection of the SARS-CoV-2 virus was performed from nasopharyngeal swabs  
167 at SYNLAB Diagnostics (Barcelona, Spain) by PCR using TaqMan™ 2019-nCoV Assay Kit according to  
168 the manufacturer's protocol (Catalog number: A47532, Thermo Fischer Scientific Inc.). Viral load was  
169 quantified from nasopharyngeal swabs at IrsiCaixa laboratory (Badalona, Spain) by PCR amplification,  
170 based on the 2019-Novel Coronavirus Real-Time RT-PCR Diagnostic Panel guidelines and protocol  
171 developed by the American Center for Disease Control and Prevention (CDC).<sup>23</sup> For absolute  
172 quantification, a standard curve was built using 1/5 serial dilutions of a SARS-CoV2 plasmid (2019-  
173 nCoV\_N\_Positive Control, catalog no. 10006625,  $2 \times 10^5$  copies/ $\mu$ L, Integrated DNA Technologies) and  
174 run in parallel to all PCR determinations.

#### 175 *Outcomes and definitions*

176 Transmission was characterized by examining the number of infected and uninfected individuals among  
177 close contacts to an index case. We defined transmission events as PCR-positivity at any time point (i.e.,  
178 days 1, 14, or at any other unscheduled PCR testing when participants referred symptoms) of a contact in  
179 the same household or workplace within the 14 days following enrollment. We defined the secondary  
180 attack rate of viral transmission as the ratio of PCR-positive individuals among close contacts, according  
181 to the WHO guidelines.

182 Development to symptomatic disease was defined as presence of at least one of the following symptoms:  
183 fever, cough, difficulty breathing, myalgia, headache, sore throat, new olfactory and taste disorder(s), or  
184 diarrhea) and a positive SARS-CoV-2 RT-PCR test. The incubation period was defined as time from first  
185 exposure to symptom onset, with later confirmation of infection by PCR.<sup>24</sup> The earliest possible exposure  
186 with the symptomatic index case was determined for each contact individually.

#### 187 Study Participants

188 We selected all eligible individuals within the original trial population for each of the three analyses  
189 conducted in the current study. As in the original trial there was no evidence of an impact of  
190 hydroxychloroquine on either transmission or development of symptomatic disease we included  
191 individuals in both arms of the trial in the current study. Firstly, all Covid-19 cases with quantitative PCR  
192 data were included in an analysis of the association between clinical and demographic features of cases  
193 and viral load. Secondly, we identified factors associated with transmission using all clusters of an index  
194 case (i.e., a Covid-19 case with at least one close contact) and their corresponding contacts for which  
195 quantitative viral load was available for the index case. Finally, we assessed the risk of developing  
196 symptomatic disease and the variation in the incubation period amongst all contacts with a positive PCR  
197 result at baseline, irrespective of available data of their index case.

#### 198 *Statistical Analysis*

199 We used log transformed viral loads which were approximately normally distributed and which also align  
200 with common reporting norms. The relationship between characteristics of cases and viral load was  
201 assessed using linear regression considering age (in years), sex, the number of days from reported  
202 symptom onset and the presence or absence of five key clinical features namely fever, cough, shortness of  
203 breath or rhinitis and anosmia. To identify risk factors for transmission, we used logistic regression model  
204 for the risk of transmission utilizing a random-effect model to allow for within cluster variation in the risk  
205 of transmission. Factors with potential influence on the risk of transmission included characteristics of the  
206 potential transmitter (i.e., age, sex, viral load, and the presence or absence of respiratory symptoms) and  
207 contacts (i.e., age, sex, and the type of contact they had with the index case). Finally the risk of  
208 developing symptomatic Covid-19 was assessed by fitting a cox-regression model considering the age (in  
209 years) and sex of the individual, the presence or absence of cardiovascular disease, chronic respiratory  
210 disease and the initial viral load in relation to the time to development of symptomatic disease. Data at 14  
211 days after the first study visit were censored, in line with the follow-up conducted in the original trial. All  
212 analyses were conducted in R version 4.0.

#### 213 *Role of the funding source*

214 The funder of the study had no role in the study design, data collection, data analysis, data interpretation,  
215 or writing of the report. All authors had full access to all the data in the study and had final responsibility  
216 for the decision to submit for publication.

217



## 218 **RESULTS**

### 219 *Sample characteristics*

220 During the investigation period, we identified 314 cases in whom the viral load was tested. Overall, 220  
221 (70.0%) were female and the median age was 41 (IQR 31-52). Of them, 282 had at least one close  
222 contact, resulting in the corresponding clusters, with a total of 753 contacts. Clusters had a median of 2  
223 contacts (IQR 1-3) and a maximum of 19 contacts. Most index cases of the clusters were female (n= 202,  
224 71.6%), with an average age of 42 years (SD 13 years) (Table 1).

### 225 *Index case viral load*

226 The first study visit was performed a median of 4 days (IQR 3 -5) after symptom onset. At the first study  
227 visit, the mean viral load amongst Covid-19 cases was  $10^8$  ( $10^{1.8}$ ). In multivariable linear regression the  
228 viral load amongst cases was higher in individuals who reported fever (Table 2) and negatively associated  
229 with the presence of anosmia but there was no association between the age or sex of the Covid-19 case  
230 nor the presence of reported dyspnea or cough. As anticipated viral load was negatively associated with  
231 the number of days since symptom onset.

### 232 *Cluster-level transmission*

233 For our risk factor analysis on SARS-CoV-2 transmission we used linked case and contact data of 282  
234 clusters with 753 contacts. At the cluster level, 90 (33.3%) of the 282 clusters had at least one  
235 transmission event, with a highly skewed distribution of the number of transmission events per cluster  
236 (Figure 1A). The first visit for contacts took place a median of 5 days (IQR 4-7 days) after their first  
237 possible exposure to the index case. A total of 125 (16.6%) of 753 contacts had a PCR positive result over  
238 the study period. The proportion of contacts who tested positive for SARS-CoV-2 within a cluster  
239 (secondary attack rate) progressively increased with the viral load of the index case: from 12% where the  
240 index case had a viral load of  $<10^6$  copies/mL to 24% where the index case had a viral load  $>10^9$   
241 copies/mL (Figure 1B). According to the multivariate analysis, the viral load of the index case was  
242 strongly associated with the risk of onward transmission (OR per  $\log_{10}$  increase in VL 1.3; 95% CI 1.1-  
243 1.6) (Table 3). Ninety percent (114/125) of transmission events had an index case viral load of 5.1  $\log_{10}$   
244 copies/ml or more, and 50% (61/125) had a viral load of 8.8  $\log_{10}$  copies/ml or more. Other factors  
245 associated with an increased risk of transmission were household contact (OR 2.7, 95% 1.4-5.06) and age  
246 of the contact (OR 1.02, 95% 1.01-1.04). There was no association of risk of transmission with reported  
247 mask usage by contacts, with the age or gender of the index case nor with the presence of respiratory  
248 symptoms in the index case at the initial study visit (Table 3).

249 We did not find any evidence of an association between the viral load of the index cases and the first viral  
250 load of incident positive results amongst contacts ( $p = 0.1$ , **Supplementary Appendix**) and this remained  
251 true when adjusting for both the day of illness on which the index cases baseline viral load was measured  
252 and the number of days until the contact was enrolled ( $p = 0.18$ ). Also, after excluding contacts who were  
253 PCR positive at the first study visit, we found no association between the viral load of the index case and  
254 the time to onset of incident SARS-CoV-2 infection (HR 1.01 95% CI 0.83-1.23).

### 255 *Risk factor for Covid-19 disease among PCR+ contacts*

256 Overall, 449 contacts had a positive PCR result at first visit regardless of availability on viral load data of  
257 their index case ( $n=125$ ) or not ( $n=324$ ). Twenty-eight (6.3%) of 449 contacts had symptoms at the first  
258 visit and 181 (40.3%) developed symptomatic Covid-19 within the follow-up period. The multivariable  
259 cox-regression analysis, after adjusting for age and sex, revealed that increasing viral load levels of the  
260 contact at day 1 were associated with an increased risk of developing symptomatic disease. The risk of  
261 symptomatic disease was approximately 25% amongst individuals with an initial viral load of  $<10^7$   
262 copies/mL compared to a more than 60% amongst those with an initial viral load of  $>10^9$  (HR per  $\log_{10}$   
263 increase in VL 1.12; 95% CI 1.05 – 1.2;  $p = 0.0006$ ) (**Figure 2A**). In the multivariable analysis there was  
264 no association between sex or age of individuals nor the presence of diabetes, cardiovascular or  
265 respiratory disease and the risk or time to developing symptomatic Covid-19.

266 The median time from exposure to symptom onset was 7 days (IQR 5 – 9). The time to onset of  
267 symptomatic disease decreased from a median of 7 days (IQR 5 – 10) for individuals with an initial viral  
268 load  $<10^7$  copies/mL to 6 days (IQR 4 – 8) and 5 days (IQR 3 – 8) for individuals with an initial viral load  
269 of  $10^7$ - $10^9$  and  $>10^9$  copies/mL, respectively (**Figure 2B**). Overall, 110/181 (60.8%) of participants who  
270 developed symptoms did so before day 8, 45/181 (24.9%) between days 8-10, and 22/181 (12.2%)  
271 between days 11-14.

272

## 273 **DISCUSSION**

274 In our study, we found that increasing viral load values in nasopharyngeal swabs of Covid-19 cases were  
275 associated with the greater risk of transmission measured by SARS-CoV-2 PCR positivity among  
276 contacts and also a higher risk of transmission in household environment compared to other indoor  
277 situations. In addition, we found that higher viral loads in swabs of asymptomatic contacts were  
278 associated with higher risk of developing symptomatic Covid-19 and have shorter incubation periods than  
279 those with a lower viral load. Relationships between viral load and infectivity have been described for  
280 other respiratory viruses and our study confirms the same is true for SARS-CoV-2.

281 To our knowledge this is the largest study that evaluates the relationship of viral load in Covid-19 cases  
282 and risk of transmission. In our cohort, a high proportion (67%) of index cases did not cause secondary  
283 infections. However, we identified 90 (33%) clusters with transmission events and the multivariate  
284 analysis revealed that clusters centered on index cases with high viral load were significantly more likely  
285 to result in transmission. Secondary attack rate was under 12% when the index case viral load was  $<10^6$   
286 copies/ml compared to more than 20% amongst clusters with the highest viral loads. In line with previous  
287 analyses of case-contact clusters,<sup>9,12,14</sup> we also found that household exposure to an index case was  
288 associated with a higher risk of transmission than other types of contact, presumably reflecting duration  
289 and proximity of exposure. Age of the contact was also identified in our multivariate analysis as a  
290 significant—albeit modest—determinant of transmission. This factor has shown uneven influence across  
291 results reported elsewhere, but seems to play a secondary role among adults.<sup>13,14</sup> Finally, unlike previous  
292 analyses that reported a relationship between coughing and transmission,<sup>13</sup> we did not find any  
293 association. This finding suggests that the absence of cough does not preclude significant onward  
294 transmission, particularly if the viral load is high. Taken together, our results indicate that the viral load,  
295 rather than symptoms, may be the predominant driver of transmission.

296 Importantly, we report that high viral load shortly after exposure in asymptomatic contacts was strongly  
297 associated with the risk of developing symptomatic Covid-19 disease. We found an approximately 25%  
298 chance of developing symptomatic disease amongst individuals with an initial viral load  $<10^7$  copies/mL  
299 compared to a more than 60% chance amongst individuals with a viral load  $>10^9$ . These data may provide  
300 rationale for risk stratification for developing illness. Moreover, the initial viral load significantly shifted  
301 the incubation time, which ranged from 5 days in participants with a high viral load to 7 days in  
302 participants with a low viral load. Our study is the first analysis of prospective data that investigates the  
303 association between initial viral load and the incubation time.

304 The study has several limitations. First, asymptomatic people were not enrolled as index cases, affecting  
305 our ability to fully characterize all types of transmission chain. Second, we did not find any evidence of  
306 decreased risk of transmission in individuals who reported mask use. While this finding collides with the  
307 evidence reported elsewhere,<sup>8</sup> we did not have fine-grained data on type of mask (surgical vs FFP2), use  
308 of other measures of PPE or other infection control practices, thus limiting our ability to make clear  
309 inferences about the impact of PPE on transmission risk. Mask usage is likely correlated with type of  
310 exposure which might further confound associations but we did not note any association between mask  
311 use and risk either in our unadjusted analysis (Table 3) or in a multivariable model excluding type of  
312 exposure (data not shown). Third, we used time to symptom onset (with later confirmation of infection)  
313 rather than time to positive PCR test based on serial testing. Nonetheless, accurate calculation of the

314 incubation period was feasible because of the prospective nature of the study, accurate identification of  
315 exposure by face-to-face interview, and intensive active and passive monitoring of exposed contacts. We  
316 followed participants over 14-day periods, thus incubation periods beyond 14 days may not have been  
317 detected. Within each cluster we cannot be completely certain about the directionality of transmission, but  
318 our inclusion criteria including the absence of Covid-19 like symptoms in the 2 weeks preceding  
319 enrolment is consistent with transmission from a case to a contact. We also cannot exclude that some  
320 individuals may have been infected by individuals outside of study clusters, but as per national guidelines  
321 all contacts were quarantined after exposure to index cases reducing the chance of transmission from  
322 elsewhere. Samples were available from index cases a median of four days after symptom onset and the  
323 initial sample in contacts was taken on average 5 days after exposure which may limit our ability to detect  
324 associations with peak viral load. Despite this we still demonstrate clear dose effects in relation to both  
325 risk of transmission and time to symptom onset. Finally, our study population is reflective of the trial  
326 from which the study sample is drawn and is therefore biased towards female participants, few  
327 comorbidities and predominantly mild-moderate infection and further data is needed on the risk of  
328 transmission in other populations.

329 In summary, our results provide evidence regarding the determinants of SARS-CoV-2 transmission,  
330 particularly on the role of the viral load. The higher risk of transmission among individuals with higher  
331 viral loads adds to current evidence and encourages assessing viral load in cases with a larger number of  
332 close contacts. When a case with high viral load is identified, implementation of reinforced contact  
333 tracing measures and quarantines, may be critical to reduce onward transmission. Similarly, our results  
334 regarding the risk and expected time to developing symptomatic Covid-19 encourage risk stratification of  
335 newly diagnosed SARS-CoV-2 infections based on the initial viral load.

336

337

## 338 **CONTRIBUTORS**

339 MM, DO, OM accessed and verified the data. MM, DO, ChR, OM conceived of the study. MM  
340 performed the analysis. PM, AA, MCM, MU, MVM, CGB, NP, JA, BC, OM led the RCT from which  
341 study data is derived. MM and OM wrote the first draft of the manuscript. All authors gave critical input  
342 into interpretation and revised the manuscript.

343

## 344 CONFLICTS OF INTEREST

345 We declare no conflicts of interest

346

## 347 ACKNOWLEDGMENTS

348 The authors would like to thank Gerard Carot-Sans for providing medical writing support during the  
349 preparation of the manuscript.

## 350 REFERENCES

- 351 1 La Rosa G, Bonadonna L, Lucentini L, Kenmoe S, Suffredini E. Coronavirus in water environments:  
352 Occurrence, persistence and concentration methods - A scoping review. *Water Res* 2020; **179**: 115899.
- 353 2 Umakanthan S, Sahu P, Ranade A V., *et al.* Origin, transmission, diagnosis and management of  
354 coronavirus disease 2019 (COVID-19). *Postgrad Med J* 2020; **0**: 1–6.
- 355 3 Leclerc QJ, Fuller NM, Knight LE, Funk S, Knight GM. What settings have been linked to SARS-  
356 CoV-2 transmission clusters? *Wellcome Open Res* 2020; **5**: 83.
- 357 4 Qian H, Miao T, LIU L, Zheng X, Luo D, Li Y. Indoor transmission of SARS-CoV-2. *medRxiv* 2020; :  
358 2020.04.04.20053058.
- 359 5 Hamner L, Dubbel P, Capron I, *et al.* High SARS-CoV-2 Attack Rate Following Exposure at a Choir  
360 Practice — Skagit County, Washington, March 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**: 606–  
361 10.
- 362 6 Park SY, Kim YM, Yi S, *et al.* Coronavirus disease outbreak in call center, South Korea. *Emerg Infect*  
363 *Dis* 2020; **26**: 1666–70.
- 364 7 Muñoz MA, López-Grau M. Lessons learned from the approach to the COVID-19 pandemic in urban  
365 primary health care centres in Barcelona, Spain. *Eur J Gen Pract* 2020; **26**: 106–7.
- 366 8 Chu DK, Akl EA, Duda S, *et al.* Physical distancing, face masks, and eye protection to prevent person-  
367 to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The*  
368 *Lancet* 2020; **395**: 1973–87.
- 369 9 Bi Q, Wu Y, Mei S, *et al.* Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their  
370 close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 2020; **20**: 911–9.
- 371 10 Wölfel R, Corman VM, Guggemos W, *et al.* Virological assessment of hospitalized patients with  
372 COVID-2019. *Nature* 2020; **581**: 465–9.
- 373 11 La Scola B, Le Bideau M, Andreani J, *et al.* Viral RNA load as determined by cell culture as a  
374 management tool for discharge of SARS-CoV-2 patients from infectious disease wards. *Eur J Clin*  
375 *Microbiol Infect Dis* 2020; **39**: 1059–61.

- 376 12 Böhmer MM, Buchholz U, Corman VM, *et al.* Investigation of a COVID-19 outbreak in Germany  
377 resulting from a single travel-associated primary case: a case series. *Lancet Infect Dis* 2020; **20**: 920–8.
- 378 13 Wu J, Huang Y, Tu C, *et al.* Household Transmission of SARS-CoV-2, Zhuhai, China, 2020. *Clin*  
379 *Infect Dis* 2020; published online May 11. DOI:10.1093/cid/ciaa557.
- 380 14 Cheng HY, Jian SW, Liu DP, Ng TC, Huang WT, Lin HH. Contact Tracing Assessment of COVID-19  
381 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods before and after Symptom  
382 Onset. *JAMA Intern Med* 2020; **180**: 1156–63.
- 383 15 Huang L, Zhang X, Zhang X, *et al.* Rapid asymptomatic transmission of COVID-19 during the  
384 incubation period demonstrating strong infectivity in a cluster of youngsters aged 16-23 years outside  
385 Wuhan and characteristics of young patients with COVID-19: A prospective contact-tracing study. *J*  
386 *Infect* 2020; **80**: e1–13.
- 387 16 Liu T, Gong D, Xiao J, *et al.* Cluster infections play important roles in the rapid evolution of COVID-  
388 19 transmission: a systematic review. *Int J Infect Dis* 2020; **99**: 374.
- 389 17 Buitrago-Garcia D, Egli-Gany D, Counotte MJ, *et al.* Occurrence and transmission potential of  
390 asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-  
391 analysis. *PLOS Med* 2020; **17**: e1003346.
- 392 18 Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019- nCoV)  
393 infections among travellers from Wuhan, China, 20 28 January 2020. *Eurosurveillance* 2020; **25**: 20–  
394 8.
- 395 19 Li Q, Guan X, Wu P, *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-  
396 infected pneumonia. *N Engl J Med* 2020; **382**: 1199–207.
- 397 20 Leung C. The difference in the incubation period of 2019 novel coronavirus (SARS-CoV-2) infection  
398 between travelers to Hubei and nontravelers: The need for a longer quarantine period. *Infect Control*  
399 *Hosp Epidemiol* 2020; **41**: 594–6.
- 400 21 Mitja O, Ubals M, Corbacho M, *et al.* A Cluster-Randomized Trial of Hydroxychloroquine as  
401 Prevention of Covid-19 Transmission and Disease. *medRxiv* 2020; : 2020.07.20.20157651.
- 402 22 Catalan Ministry of Health. Catalan epidemiological surveillance system.  
403 [http://salutpublica.gencat.cat/ca/ambits/vigilancia\\_salut\\_publica/](http://salutpublica.gencat.cat/ca/ambits/vigilancia_salut_publica/) (accessed March 28, 2020).
- 404 23 Centers for Disease Control and Prevention (CDC). CDC 2019-Novel Coronavirus (2019-nCoV) Real-  
405 Time RT-PCR Diagnostic Panel. Cat. 2019-NCovEUA-01. 2020.  
406 <https://www.fda.gov/media/134922/download> (accessed May 21, 2020).
- 407 24 World Health Organization. The First Few X (FFX) Cases and contact investigation protocol for 2019-  
408 novel coronavirus (2019-nCoV) infection, version 2. [https://www.who.int/publications/i/item/the-first-few-x-\(ffx\)-cases-and-contact-investigation-protocol-for-2019-novel-coronavirus-\(2019-ncov\)-infection](https://www.who.int/publications/i/item/the-first-few-x-(ffx)-cases-and-contact-investigation-protocol-for-2019-novel-coronavirus-(2019-ncov)-infection)  
409 (accessed Sept 21, 2020).  
410

412 **Tables**

413 **Table 1:**

414 **Baseline Characteristics of linked transmission clusters**

Variable		Value
Cluster Size	Median (IQR)	2 (1-3)
Index Case Age	Years – Mean (SD)	42 (13)
Index Case Sex	Female (%)	202 (64.%)
Index Case Log Viral Load	Mean (SD)	8 (1.8)
Contacts Age	Years – Mean (SD)	42 (15)
Contacts Gender	Female (%)	385 (51.1%)
	Male (%)	305 (40.5%)
	Missing (%)	63 (8.4%)
Baseline PCR of Contact Case	Positive	93 (14.2%)
Contact	HCW	254 (33.7%)
	Household	382 (50.7%)
	Nursing Home	21 (2.8%)
	Unknown	96 (12.7%)

415

416

417 **Table 2: Univariable and multivariable linear regression of association between Index case**  
 418 **variables and log<sub>10</sub> viral load**

Characteristic		Log <sub>10</sub> Viral Load/ml	Unadjusted $\beta$ coefficient (95% CI)	p	Adjusted $\beta$ coefficient (95% CI)	p
Case Age		N/A	0.002 (-0.02 – 0.02)	0.78	0.008 (-0.01 – 0.02)	0.38
Case Sex	Male	8.15 (7.54 – 8.77)	Reference		Reference	
	Female	8.04 (7.47 – 8.6)	-0.238 (-0.72 – 2.4)	0.33	-0.22 (-0.61 – 0.34)	0.59
Days from Symptom Onset		NA	-0.17 (-0.26 – 0.0.8)	0.0002	-0.16 (-0.24 – 0.07)	0.0004
Cough	Absent	7.82 (7.24 – 8.41)	Reference		Reference	
	Present	8.37 (7.78 – 8.95)	0.66 (0.22 – 1.1)	0.003	0.41 (-0.02 – 0.84)	0.06
Dyspnea	Absent	7.97 (7.5-8.43)	Reference		Reference	
	Present	8.22 (7.45-8.99)	0.27 (-0.40 – 0.94)	0.42	0.28 (-0.35 – 0.92)	0.38
Fever	Absent	7.77 (7.16 – 8.38)	Reference		Reference	
	Present	8.42 (7.86-8.98)	0.80 (0.36 – 1.24)	0.0004	0.43 (0.00 – 0.87)	0.05
Anosmia	Absent	8.32 (7.76 – 8.88)	Reference		Reference	
	Present	7.87 (7.25-8.49)	-0.57 (-1.0 - -0.09)	0.02	-0.54 (-1.0 – 0.09)	0.02
Rhinitis	Absent	7.60 (7.23 – 7.98)	Reference		Reference	
	Present	8.59 (7.65-9.52)	0.88 (-0.05 – 1.82)	0.06	0.77 (-0.11 – 1.66)	0.09

419

420



421

422 **Table 3: Risk factors for transmission of SARS-CoV-2**

		<b>Unadjusted Odds Ratio</b>	<b>Confidence Interval</b>	<b>p</b>	<b>Adjusted Odds Ratio</b>	<b>Confidence Interval</b>	<b>p</b>
Index case age (per year)		1.02	0.99-1.05	0.07	1	0.99-1.03	0.46
Female Index Case		0.74	0.4-1.36	0.33	0.71	0.37-1.39	0.32
Index Case Viral Load (per Log <sub>10</sub> change)		1.27	1.09-1.48	<0.01	1.29	1.1-1.5	0.001
Index Case Cough		1.0	0.55-1.82	0.99	1.13	0.64 – 2.0	0.66
Index Case Dyspnea		0.80	0.31-2.07	0.64	0.75	0.30 – 1.89	0.55
Index Case Rhinitis		1.46	0.46-4.63	0.52	1.31	0.42-4.11	0.64
Age of Contact		1.03	1.01-1.05	<0.01	1.02	1.01 – 1.04	0.0008
Female Contact		0.93	0.58-1.49	0.77	1.33	0.79 – 2.23	0.28
Mask Use	Never	1 (Reference Group)	N/A	N/A	1 (Reference Group)	N/A	N/A
	Always	0.93	0.47 – 1.83	0.84	1.55	0.76 – 3.16	0.23
	Unknown	1.18	0.59 – 2.36	0.47	1.49	0.74-3.01	0.0.26
Contact Type	Healthcare Work	1 (Reference Group)	N/A	N/A	1 (Reference Group)	N/A	N/A
	Household	3.07	1.68-5.62	<0.01	3.0	1.59 – 5.65	0.0006
	Nursing Home	1.75	0.19 -16.01	0.62	1.90	0.30 – 11.91	0.49
	Other	0.32	0.03-3.05	0.32	1.19	0.10 – 14.31	0.89

423

424

425 **Figure 1: Transmission in a cluster**

426

427 (A) Number of secondary cases per cluster. (B) Relationship between viral load of the index case and the  
428 proportion of contacts developing Covid-19. Numbers 18/149 in group  $10^4$ - $10^5$  RNA copies/ml; 30/2012  
429 in group  $10^6$ - $10^7$ ; 59/298 in group  $10^8$ - $10^9$ ; 17/71 in group  $\geq 10^{10}$ .

430

431

432 **Figure 2. Risk of developing symptomatic Covid-19 according to characteristics of the contact at**  
433 **enrolment.**

434

435 (A) probability of symptomatic disease by viral load. (B) time to symptomatic disease by viral load.

436

437

438 **Data Sharing Statement**

439

440 Marks M, Millat-Martinez P, Ouchi d et al. Transmission of Covid-19 in 282 clusters in  
441 Catalonia, Spain: a cohort study

442

443

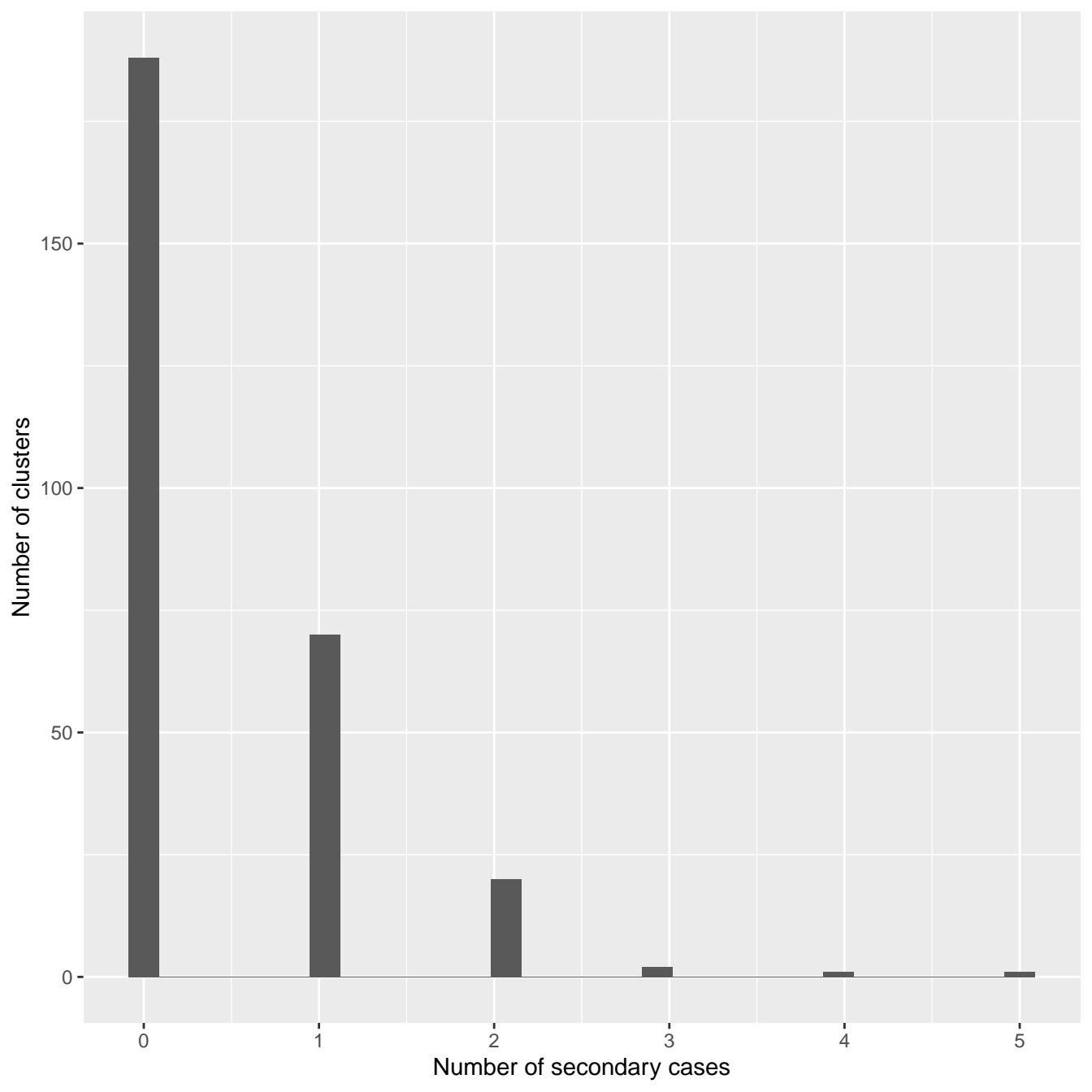
444

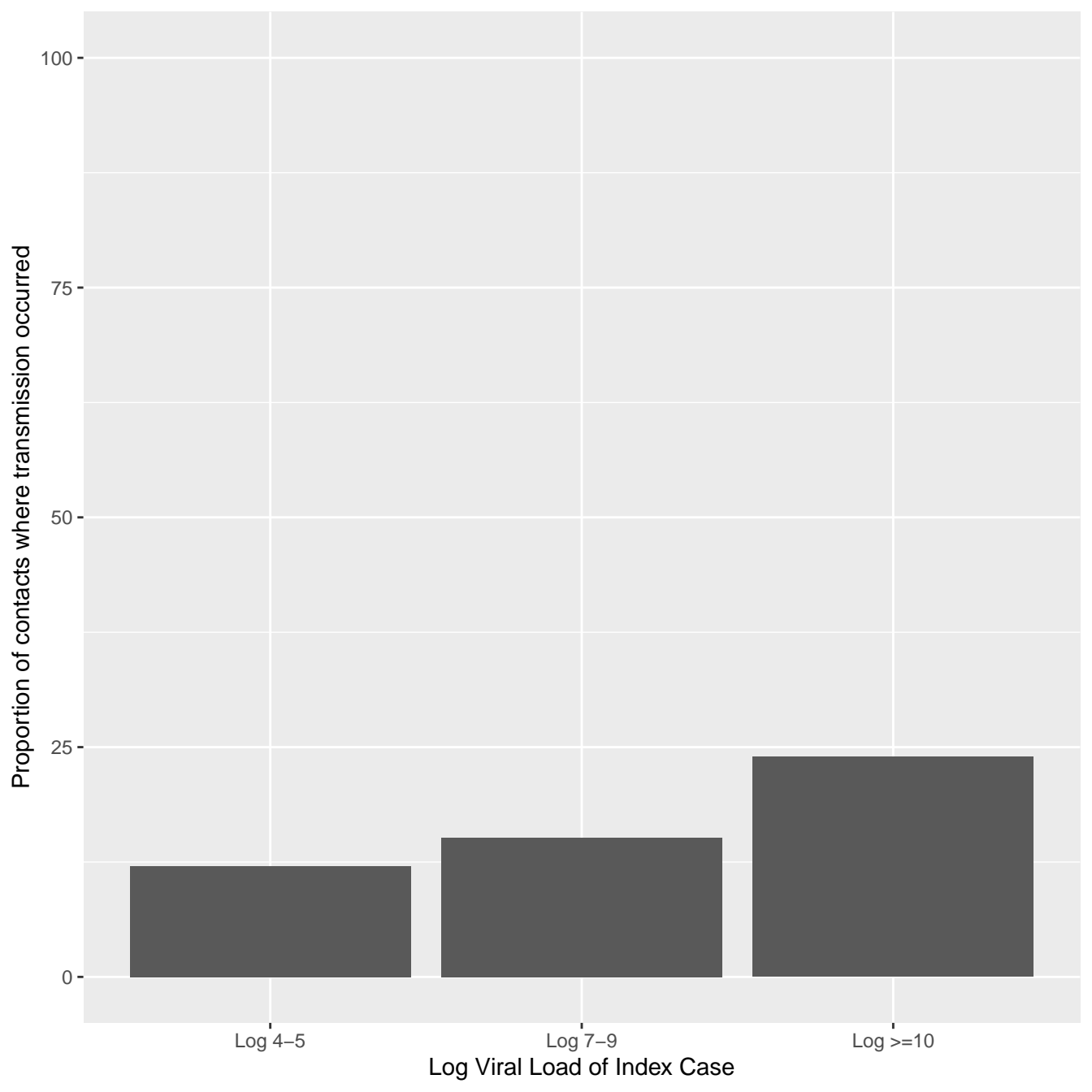
Question	Authors' Response
Will the data collected for your study be made available to others?	Yes
Which data?	Complete de-identified patient data set
Will any supporting documents be available?	Study protocol
How or where can the data be obtained?	mcorbacho@flsida.org
When will data availability begin?	15/12/2020
When will data availability end?	15/12/2021
To whom will data be available?	Researchers whose proposed use of the data has been approved
For what type of analysis or purpose?	For any purpose
By what mechanism?	By email
Any other restrictions?	—
Additional information	—

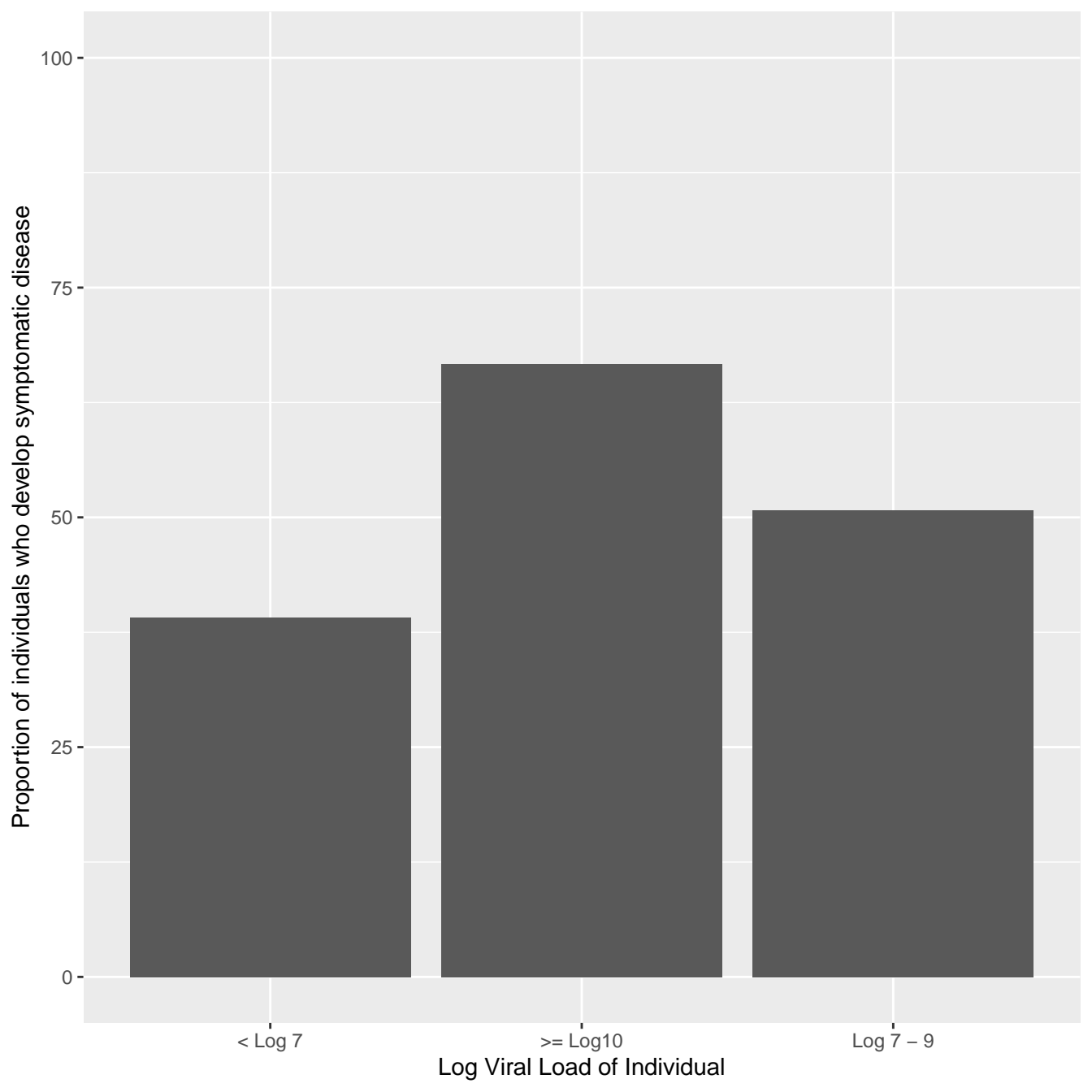
445

446

447







100

75

50

25

0

Proportion of individuals who develop symptomatic disease

< Log 7

$\geq$  Log 10

Log 7 - 9

Log Viral Load of Individual

